UNIVERSITY
OF TASMANIA

Chlorinated Auxin in the Fabaceae: Distribution, Evolutionary Origin and Genetic Aspects

By

Hong Kiat Lam, BBiotech (Hons)

School of Biological Sciences

Submitted in fulfilment of the requirements of the degree of Doctor of Philosophy

University of Tasmania

June 2017

Declaration of Originality

I hereby declare that this thesis contains no material which has been accepted for a degree or diploma by the University or any other institution, except by way of background information and duly acknowledged in the thesis, and to the best of my knowledge and belief no material previously published or written by another person except where due acknowledgement is made in the text of the thesis, nor does the thesis contain any material that infringes copyright.

Access statement

This thesis may be made available for loan. Copying and communication of any part of this thesis is prohibited for two years from the date this statement was signed; after that time limited copying and communication is permitted in accordance with the Copyright Act 1968.

_____

Hong Kiat Lam
School of Biological Sciences
University of Tasmania
June 2017

Statement of Co-Authorship

The following people and institutions contributed to the publication of work undertaken as part of this thesis:

Name and School = **Hong Kiat Lam, School of Biological Sciences**
Name and institution = **Hong Kiat Lam, University of Tasmania**
Name and institution = **Scott A.M. McAdam, University of Tasmania**
Name and institution = **Erin L. McAdam, University of Tasmania**
Name and institution = **John J. Ross, University of Tasmania**

Author details and their roles:

Paper 1, <Evidence that chlorinated auxin is restricted to the Fabaceae but not to the Fabeae>:

Located in chapter 2

Hong Kiat Lam conducted the experiments and wrote part of the article; Scott McAdam obtained the plant material, supervised the phylogenetic analysis, and wrote part of the article; Erin McAdam supervised the hormone analyses and complemented the writing; John Ross initiated the project and wrote part of the article.

Paper 2, <The single evolutionary origin of chlorinated auxin provides a phylogenetically Informative trait in the Fabaceae>:

Located in chapter 3

Hong Kiat Lam conducted the experiments and wrote part of the article; Scott McAdam initiated the project, obtained the plant material, supervised the phylogenetic analysis, and wrote part of the article; Erin McAdam supervised the hormone analyses and complemented the writing; John Ross wrote part of the article.

We the undersigned agree with the above stated "proportion of work undertaken" for each of the above published (or submitted) peer-reviewed manuscripts contributing to this thesis:

Signed: _____ _____

John J. Ross
Supervisor
School Of Biological Sciences
University fo Tasmania

Mark Hunt
Head of School
School of Biological Sciences

University of Tasmania

Date: 29th June 2017 30th June 2017

Abstract

The ubiquitous presence in plants of a group of hormones called auxins, which includes indole-3-acetic acid (IAA), 4-chloroindole-3-acetic acid (4-Cl-IAA), indole-3-butyric acid (IBA) and phenylacetic acid (PAA), reflects their pivotal roles in mediating various aspects of plant growth and morphogenesis. However, the level of 4-Cl-IAA greatly exceeds the level of primary auxin, IAA in maturing pea (*Pisum sativum*) seeds. This highly potent chlorinated auxin, 4-Cl-IAA, and its precursor, chlorinated tryptophan (4-Cl-Trp) are mainly found in seeds but were once thought to be restricted to species of the tribe *Fabeae* within the Fabaceae. One curious exception was *Pinus sylvestris*, (Scots pine), which was also reported to contain chlorinated auxin in seeds.

The action and activity of halogenating enzymes, however, remains unknown in plants. This thesis reports a comprehensive study on the distribution of the chlorinated auxin in the Fabaceae, by screening key species in the Fabaceae and the genus *Pinus*. For the first time, 4-Cl-IAA was detected in the reproductive structures of species in the genera *Trifolium*, *Melilotus*, *Trigonella*, *Medicago* and *Ononis*, all beyond the tribe *Fabeae*. Scots pine and four other *Pinus* species, including *P. flexilis*, *P. pinea*, *P. radiata* and *P. parviflora*, were retested but none contained detectable levels of 4-Cl-IAA. This supports that 4-Cl-IAA is unique to the Fabaceae in all plant species. The notion that chlorinated auxin is restricted to reproductive structures was supported by the evidence of much higher levels of the hormone in the seeds than vegetative tissues in broad bean (*Vicia faba*) in this study.

No detectable level of 4-Cl-IAA was found in the seeds of the cultivated species *Cicer arietinum* as well as its wild relative, *C. echinospermum*, both from the tribe *Cicerae*, immediately basal to the tribe *Fabeae* and *Trifolieae*. Five species (*Galega officinalis*, *Galega bicolor*, *Parochetus communis*, *Astragalus propinquus* and *A. sinicus*) from tribes and clades basal or sister to the tribe *Cicerae* were also found to not contain detectable level of 4-Cl-IAA. The absence of 4-Cl-IAA in these species suggests a single evolutionary origin of 4-Cl-IAA, which can be used as a phylogenetically informative trait within the Fabaceae. Interestingly, the evidence indicates that some *Trigonella* species do not produce 4-Cl-IAA in their seeds. This

loss of chlorinating ability is thought to be a derived state in the genus and unique to this genus within the *Trifolieae*, shedding light on interspecific relationships within the genus. A previously published putative vanadium dependent haloperoxidase gene sequence was used to generate pea haloperoxidase mutants by Targeting Induced Local Lesions in Genomes (TILLING). No obvious phenotypes were observed among the mutants in relation to 4-Cl-IAA synthesis and hormone levels, suggesting the possibility of gene redundancy

Acknowledgements

First and foremost, I would like to express my deepest heartfelt gratitude to my supervisors John Ross, Scott McAdam, Frances Sussmilch and Jim Weller. Your enthusiasm, motivation, patience, guidance and generosity of sharing knowledge as well as the effort and time spent helping me throughout of my PhD life have made me indebted to you all. Thanks to you, my study comes to a completion, and I have gained the skills and inspiration to continue my journey as a plant scientist. I am also grateful to Sandra Davidson, Jenny Smiths, Erin McAdam, Laura Quittenden for guiding me through molecular works and hormonal analysis.

I would like to thank David Nichols and Noel Davies for their technical expertise in all things about UPLC-MS. Without their effort analysing countless samples for me, this extensive study of chlorinated auxin spanning several genera in the Fabaceae would not have been possible. A big thanks to Michelle Lang and Tracey Winterbottom for their limitless patience and effort to keep my plants healthy and happy. I am also thankful to Clancy Carver, Catherine Jones, Jodi Noble and Morgan Green for their administrative assistance at the School of Biological Sciences.

I am also grateful to the other members of the Pea Hormone Group, especially Sam Cook, Ariane Gelinas-Marion and Shelly Urquhart for their help and friendship. I also owe huge thanks to my fellow colleagues, who are also in pursuit of their PhD, Raul Ortega, Le Son and Bich Nguyen for their company and kindness tolerating my still expanding indoor forest of tropical Hoya plants, and the occasional outbreaks of fungus gnats inside our shared student office.

I would like to share my happiness of completing this journey with my beloved parents and younger brother for their love and care regardless of the physical distance. 弟、爸妈，有你们真好，我愛你们！Last, but not least, I would like to thank a bunch of awesome old and new friends plus the beloved special one who have supported me all through the years from all over the world and home.

Abbreviations

4-Cl-IAA	4-chloroindole-3-acetic acid
4-Cl-Trp	4-chlorotryptophan
BHT	butylated hydroxytoluene
dNTP	deoxynucleotide triphosphate
EDTA	ethylenediaminetetraacetic acid
EtOH	ethanol
FW	fresh weight
h	hours
HRM	high resolution melt
IAA	indole-3-acetic acid
IPyA	indole-3-pyruvic acid
m/z	mass:charge ratio
MRM	multiple reaction monitoring
MS	mass spectrometry
PCR	polymerase chain reaction
rpm	revolutions per minute
RT	retention time
TILLING	Targeting Induced Local Lesions IN Genomes
Trp	Tryptophan
UPLC	ultra-performance liquid chromatography
WT	wild-type

Table of Contents

Declaration of Originality ..... **i**

Access statement ..... **i**

Statement of Co-Authorship ..... **ii**

Abstract ..... **iii**

Acknowledgements ..... **v**

Abbreviations ..... **vi**

Table of Contents ..... **vii**

Chapter 1 General Introduction to 4-Chloroindole-acetic Acid Biosynthesis and Tryptophan Halogenating Enzymes **1**

1.1 Introduction *1*

1.1.1 Naturally occurring chlorinated auxin, 4-Cl-IAA 5

1.1.1.1 Occurrence and tissue localisation 5

1.1.1.2 Biosynthesis and biological roles 6

1.1.2 A panoply of biohalogenation 9

1.1.3 Tryptophan halogenating enzymes 13

1.2 Aims for this study *16*

Chapter 2 Distribution of Chlorinated Auxins in Plants **17**

2.1 Introduction *17*

2.2 Materials and methods *19*

2.2.1 Plant materials 19

2.2.2 Extract preparation for the detection and quantification of compounds 19

2.3 Results *21*

2.4 Discussion *38*

Chapter 3 The Single Evolutionary Origin of Chlorinated Auxin in the Fabaceae **42**

3.1 Introduction *42*

3.2 Materials and methods *43*

3.2.1 Plant materials 43

3.2.2 Extract preparation for the detection and quantification of compounds 43

3.3 Results and discussion *44*

Chapter 4 Phylogenetic Relationships in the Fabaceae: Insights from the Possible Loss of Chlorinating Ability in Some Members of the Genus *Trigonella* **56**

4.1 Introduction *56*

4.1.1 The Vicioids 58

4.1.2 *Trigonella* 60

4.1.3 Resolving taxonomic placement: choice of useful characteristics 60

4.2 Materials and methods *64*

4.2.1 Plant materials 64

4.2.2 Extract preparation for the detection and quantification of compounds 64

4.3 Results and discussion *66*

Chapter 5 Mutant-based Study of a Vanadium-dependent Haloperoxidase Gene in *Pisum sativum* 98

5.1 INTRODUCTION *98*

5.2 Materials and methods *103*

5.2.1 Sequence availability 103

5.2.2 Plant materials 104

5.2.3 Primers 104

5.2.4 Gene isolation 104

5.2.4.1 DNA extraction 104

5.2.4.2 DNA quantification 105

5.2.4.3 Polymerase Chain Reaction (PCR) 105

5.2.4.4 Visualisation of DNA 106

5.2.4.5 PCR product purification 106

5.2.4.6 Sequencing 106

5.2.5 Gene expression 106

5.2.5.1 RNA extraction 106

5.2.5.2 Quantification 106

5.2.5.3 Reverse transcription 107

5.2.5.4 Quantitative reverse transcription PCR (qRT-PCR) 107

5.2.6 TILLING 108

5.2.7 Design and use of molecular markers for genotyping 108

5.2.7.1 Derived Cleaved Amplified Polymorphic Sequence (dCAPS) markers 108

5.2.7.2 High Resolution Melt (HRM) markers 109

5.2.8 Construction of alignments and phylogenetic trees 109

5.2.9 Statistical analysis 110

5.3 RESULTS AND DISCUSSION *111*

5.3.1 Haloperoxidase or not? 125

Chapter 6 General Discussion and Conclusion 131

References 136

Appendix 1: Primer details 142

Chapter 1 General Introduction to 4-Chloroindole-acetic Acid Biosynthesis and Tryptophan Halogenating Enzymes

1.1 Introduction

Auxins are a structurally diverse class of phytohormones that essentially mediate and control a vast array of plant growth and morphogenesis in response to environmental cues. They are endogenously found in all plant phyla, algae, fungi and bacteria (Ross and Reid, 2010; Zhang and van Duijn, 2014). It has been shown that auxins are involved in the control of cell division and cell elongation (Sauer *et al.*, 2013), the regulation of fruit formation (De Jong *et al.*, 2009; Obroucheva, 2014; Pattison *et al.*, 2014), the control of senescence (Im Kim *et al.*, 2011; Basu *et al.*, 2013; Mueller-Roeber and Balazadeh, 2013), the establishment and maintenance of polarity and tropic response (Woodward and Bartel, 2005; Vanneste and Friml, 2009), and responses to pathogens (Fu and Wang, 2011; Mutka *et al.*, 2013) and abiotic stress (Wang *et al.*, 2010; Kodaira *et al.*, 2011). Research on auxins began when the presence of a 'transmitted influence' was postulated by Theophil Ciesielski and this idea was further developed by Charles and Frances Darwin (Tivendale *et al.*, 2014). The underlying substance was then isolated by Salkowski in 1885 and identified as indole-3-acetic acid (IAA) (Mano and Nemoto, 2012; Sauer *et al.*, 2013).

Two major routes to IAA, the predominant representative of auxins, are the tryptophan (Trp)-dependent and Trp-independent pathways. Four interconnected Trp-dependent IAA biosynthetic pathways have been proposed (Fig. 1.1), namely the indole-3-pyruvic acid (IPyA), indole-3-acetamide (IAM), indole-3-acetaldoxime (IAOx) and tryptamine (TAM) pathways, each named after the intermediate immediately downstream of Trp (Tivendale *et al.*, 2014). Indole-3-gylcerol phosphate (IGP) or indole is potentially the precursor in the Trp-independent pathway, although little is known about the biochemical pathway to IAA bypassing tryptophan altogether (Mano and Nemoto, 2012).

Despite being one of the oldest fields of experimental plant research, there has been much confusion and debate over *de novo* auxin biosynthesis. Due to the ubiquitous

and important status of auxins in living organisms, the incorporation of data from all species has led to a panoply of pathways to auxin in a manner of great complexity. Different pathways and strategies may be adopted by different species with the possibility of two or multiple pathways working concurrently, further adding to the complexity. This complexity has been extensively covered in many reviews [See Woodward and Bartel (2005), Zhao 2010, Korasick *et al*. (2013) and Tivendale *et al*. (2014)]. Discovery of a consensus or primary biosynthetic scheme has been elusive and challenging. Alternatively, elucidation of predominant pathway(s) in a single species or among related species may be an easier, more logical and useful approach to depict auxin biology.

In addition to the extensively studied predominant auxin, IAA, three other naturally occurring auxins have been described and detected as free acids and in conjugated forms in plants (Park *et al.*, 2010; Ludwig-Müller, 2011; Korasick *et al.*, 2013). These include 4-chloroindole-3-acetic acid (4-Cl-IAA), phenylacetic acid (PAA) and indole-3-butyric acid (IBA) (Fig. 1.2). This thesis will focus on the biosynthesis of the only chlorinated auxin, 4-Cl-IAA, and give a general overview of indole/tryptophan halogenating enzymes.

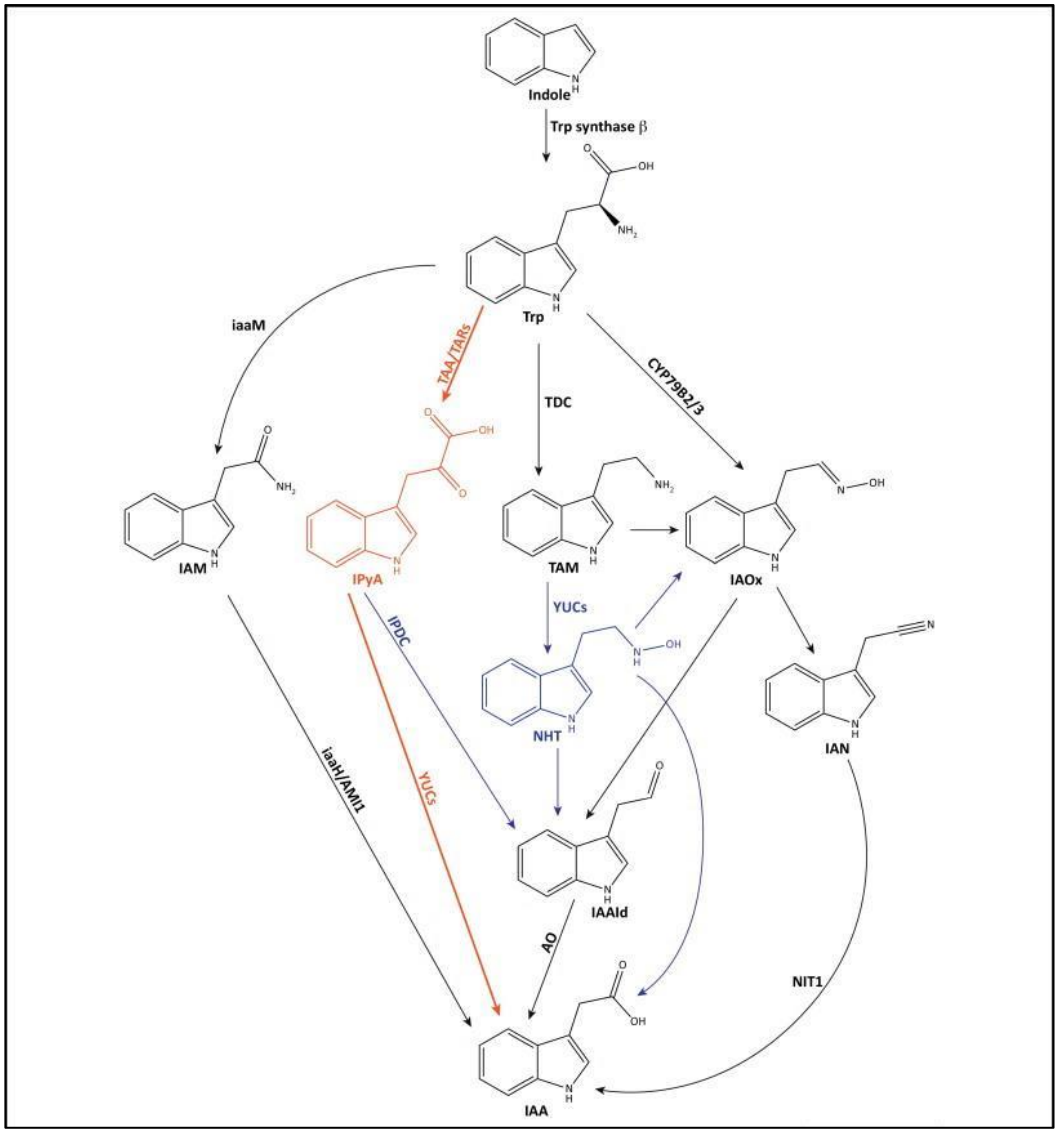
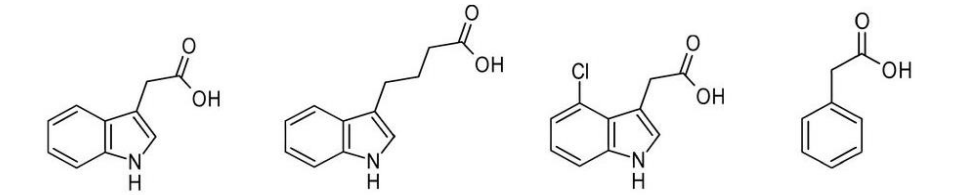


Figure 1.1 Proposed Trp-dependent auxin biosynthetic pathways. Shown in black and blue are IAA biosynthetic pathways based on data accumulated up to 2010, incorporating evidence from several species including *Arabidopsis thaliana*, *Pisum sativum, Solanum lycopersicum*, *Zea mays*, *Nicotiana tabacum*, petunia, and several species of IAA-synthesising microbes, such as *Agrobacterium tumefaciens. Pseudomonas savastanoi*, and *Enterobacter cloacae*. The two-step linear IAA biosynthetic pathway, the conversion of Trp to IPyA by TAA1/TAR enzymes and of IPyA to IAA by YUCs, based on research in *Arabidopsis thaliana*, *Pisum sativum*, and *Zea mays*, is designated in orange. Pathways suggested as unlikely, based on recent results, are shown in blue; for instance, TAM converted by YUCs to N-hydroxytryptamine (NHT). From Tivendale *et al*. 2014.



Indole-3-acetic acid (IAA) indole-3-butyric acid (IBA) 4-chloroindole-3-acetic acid (4-Cl-IAA) phenylacetic acid (PAA)

Figure 1.2 Chemical structure of the naturally occurring auxins, indole-3-acetic acid (IAA), indole-3-butyric acid (IBA), 4-chloroindole-3-acetic acid (4-Cl-IAA) and phenylacetic acid (PAA) are presented.

1.1.1 Naturally occurring chlorinated auxin, 4-Cl-IAA

1.1.1.1 Occurrence and tissue localisation

Gandar and Marumo's group first discovered 4-Cl-IAA in immature seeds of *Pisum sativum* (Ernstsen and Sandberg, 1986; Sauer *et al.*, 2013). In the Fabaceae family, specifically the *Fabeae* tribe, 4-Cl-IAA has been reported in *Pisum sativum* (pea), *Vicia faba* (broad bean), *Lens culinaris* (lentil), *Vicia sativa*, *Vicia amurensis* and four species of the genus *Lathyrus*. Endogenous 4-Cl-IAA has also been reported from both immature and mature seeds of *Pinus sylvestris* (Scots Pine) (Ernstsen and Sandberg, 1986).

4-Cl-IAA is absent in *Arabidopsis thaliana*, a key model plant. In addition, several Fabaceae species were reported to lack 4-Cl-IAA and these include *Cicer arietinum* (chickpea), *Glycine max* (soybean), *Phaseolus vulgaris* (common bean), *Vigna catjang* and *Dolichos lablab* (Katayama *et al.*, 1987). A number of agronomically cultivated plants were also subjected to feeding of radiolabeled chlorine but none of the following synthesised detectable levels of 4-Cl-IAA or its methyl ester: *Hordeum vulgare* (barley), *Avena sativa* (oat), *Triticum aestivum* (common wheat), *Secale cereal* (rye), *Zea mays* (corn), *Allium schoenoprasum* (chives), *Lepidium sativum* (garden cress), *Linum usitatissimum* (flax), *Brassica napus* (rapeseed), *Nicotiana tabacum* (tabacco) and *Helianthus anuus* (sunflower) (Engvild, 1975).

In some biological assays, 4-Cl-IAA is ten times more active than IAA within Fabaceae species (Engvild *et al.*, 1980; Reinecke, 1999; Park *et al.*, 2010; Simon and Petrášek, 2011). The levels of 4-Cl-IAA are among the highest reported for any auxin in plant tissues in the seeds of some key legumes (Tivendale *et al*., 2012; see Reinecke 1999 for review). 4-Cl-IAA is usually localised in the immature and mature seeds in those plants but it has also been reported in the leaves of *Vicia faba* (Pless *et al.*, 1984). However, particularly in pea, 4-Cl-IAA is largely restricted to both developing and mature seeds (Katayama *et al.*, 1988).

1.1.1.2 Biosynthesis and biological roles

Chlorinated auxin biosynthesis is believed to be independent of IAA itself by using 4-Cl-Trp as precursor. Recently, three *TRYPTOPHAN AMINOTRANSFERASE OF ARABIDOPSIS* (*TAA*) homologs were isolated from pea, namely *TRYPTOPHAN AMINOTRANSFERASE RLEATED1* (*PsTAR1*), *PsTAR2* and *PsTAR3* (Tivendale *et al.*, 2012). Intriguingly, TAR1 and TAR2 enzymes encoded by *PsTAR1* and *PsTAR2* genes converted tryptophan and its 4-chlorinated analog to IPyA and 4-Cl-IPyA respectively *in vitro*, which then serve as the precursors of IAA and 4-Cl-IAA in the IPyA-pathway. Gene expression patterns have indicated that *PsTAR1* plays a key role during the early seed development, whereas PsTAR2 is a key enzyme in the later stages, which is consistent with the dramatically reduced 4-Cl-IAA levels observed in *tar2* mutant, obtained by Targeting Induced Local Lesions in Genomes (TILLING) (Tivendale *et al.*, 2012). Based on several labelled-precursor feeding studies in pea seeds, parallel IPyA- and 4-Cl-IPyA pathways have been proposed. IAA itself does not become chlorinated but Trp can be chlorinated, and therefore Trp or indole could possibly be the point of chlorination but an earlier point in the biosynthesis pathway cannot be excluded (Fig. 1.3) (Tivendale *et al.*, 2012). Previous studies have ruled out the IAOx-pathway, IAM-pathway and TAM-pathway in pea seeds (Quittenden *et al.*, 2009; Tivendale *et al.*, 2010; Tivendale *et al.*, 2012).

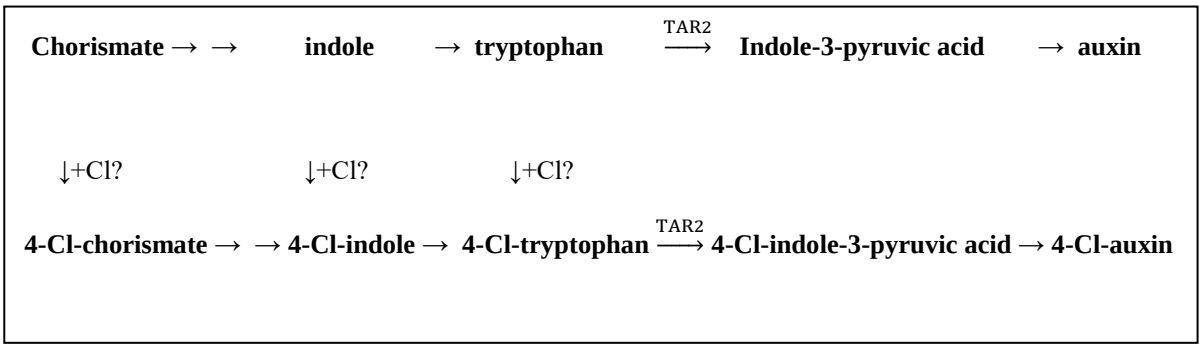


Figure 1.3 Proposed pathways of IAA and 4-Cl-IAA synthesis in pea seeds. Chlorination is known to occur at the tryptophan stage, and possibly at earlier stages as well.

Possibly due to its absence in Arabidopsis, many aspects from physiological functions to mechanisms of action of 4-Cl-IAA remain enigmatic (Simon and Petrášek, 2011). However, 4-Cl-IAA is thought to be essential for stimulating early pea seed and normal pericarp growth in pea (Reinecke *et al.*, 1995; Reinecke *et al.*, 1999; Tivendale *et al.*, 2012). For instance, exogenous application of 4-Cl-IAA has rescued and promoted the pericarp elongation of de-seeded pods (Reinecke *et al.*, 1999). Interestingly, when applied exogenously, 4-Cl-IAA is active at lower concentrations compared with IAA, which might be explained by its greater chemical stability (Sauer *et al.*, 2013). The inhibitory effect of auxin-induced ethylene on pod elongation can be hindered by 4-Cl-IAA but not IAA (Johnstone *et al.*, 2005). These differing effects of the two auxins suggest that pods may have additional or different auxin receptors, such as a TRANSPORT INHIBITOR RESPONSE1 (TIR1)-type protein which is absent in vegetative tissues, and that there might be differing affinities between 4-Cl-IAA and IAA for these auxin receptors (Ross *et al.*, 2012). Different auxin receptor proteins have been suggested to have separate functions (Parry *et al.*, 2009), and the affinity of different auxins for the Arabidopsis TIR1 protein vary (Dharmasiri *et al.*, 2005; Kepinski and Leyser, 2005). Moreover, the stimulatory effects of 4-Cl-IAA on maize coleoptile elongation reportedly correlate to its binding affinity to auxin binding sites (Rescher *et al.*, 1996), thus suggesting possible roles of 4-Cl-IAA in fast non-transcriptional responses connected with proton secretion and the maintenance of membrane potential in coleoptiles (Karcz and Burdach, 2002) and protoplast swelling (Steffens and Lüthen, 2000).

Further investigations are needed to unravel the point of halogenation and the responsible halogenase gene that incorporates the chlorine atom to the 4-position of the indole skeleton. In this study, a TILLING approach is conducted in an attempt to obtain a null mutation in a halogenase gene, so that phenotypes of relevant mutants can be characterised. Identifying the relevant genes is deemed valuable for underpinning a more comprehensive understanding of the biosynthesis, metabolism and transport of halogenated auxin in plants.

1.1.2 A panoply of biohalogenation

Over 4700 organohalogens have been isolated (Gribble, 2010) with the vast majority sourced from living organisms (Neumann *et al.*, 2008; Wagner *et al.*, 2009), including bacteria, fungi, marine algae and plants (Gribble, 1998; van Pée and Ludwig-Müller, 2002). The presence of halogen substituents, such as chlorine or bromine, iodine or fluorine, in certain positions often profoundly influences their biological activity (Reinecke, 1999; Neumann *et al.*, 2008). In nature, iodination and fluorination are rare while chlorination predominates over bromination (Murphy, 2003; Neumann *et al.*, 2008). This relatively large representation of chlorinated and brominated metabolites may reflect the abundance of chloride and bromide ions in the both marine and terrestrial environments (Vaillancourt *et al.*, 2006). A plethora of carbon-halogen bond-containing compounds ranges from peptides, polyketides, indoles, terpenes, acetogenins and phenols to volatile halogenated hydrocarbons such as bromoform, chloroform and dibromomethane (Butler and Sandy, 2009). This diverse group of halometabolites has intrigued scientists with their biological activities and functions including, but not limited to, potent anticancer, antifungal, antiviral, antibacterial, anti-inflammatory, and plant growth regulatory properties.

The prototypical heme (iron-containing porphyrin) chloroperoxiodase (CPO) was the first halogenating enzyme to be isolated. This enzyme was characterised from the caldariomycin-producing terrestrial fungus *Caldariomyces fumago* (Shaw and Hager, 1959). Since then, 6 types of halogenating enzymes have been discovered and they are other heme iron-dependent haloperoxidases, vanadium-dependent haloperoxidases, non-heme iron halogenases, flavin-dependent halogenases, and S-adenosyl-L-methionine (SAM)-dependent halogenases as well as cofactor free haloperoxidases (See Fig. 1.4 for general halogenating reactions) (Butler and Sandy, 2009; Wagner *et al.*, 2009) (Weichold *et al.*, 2016; Xu and Wang, 2016; Agarwal *et al.*, 2017).

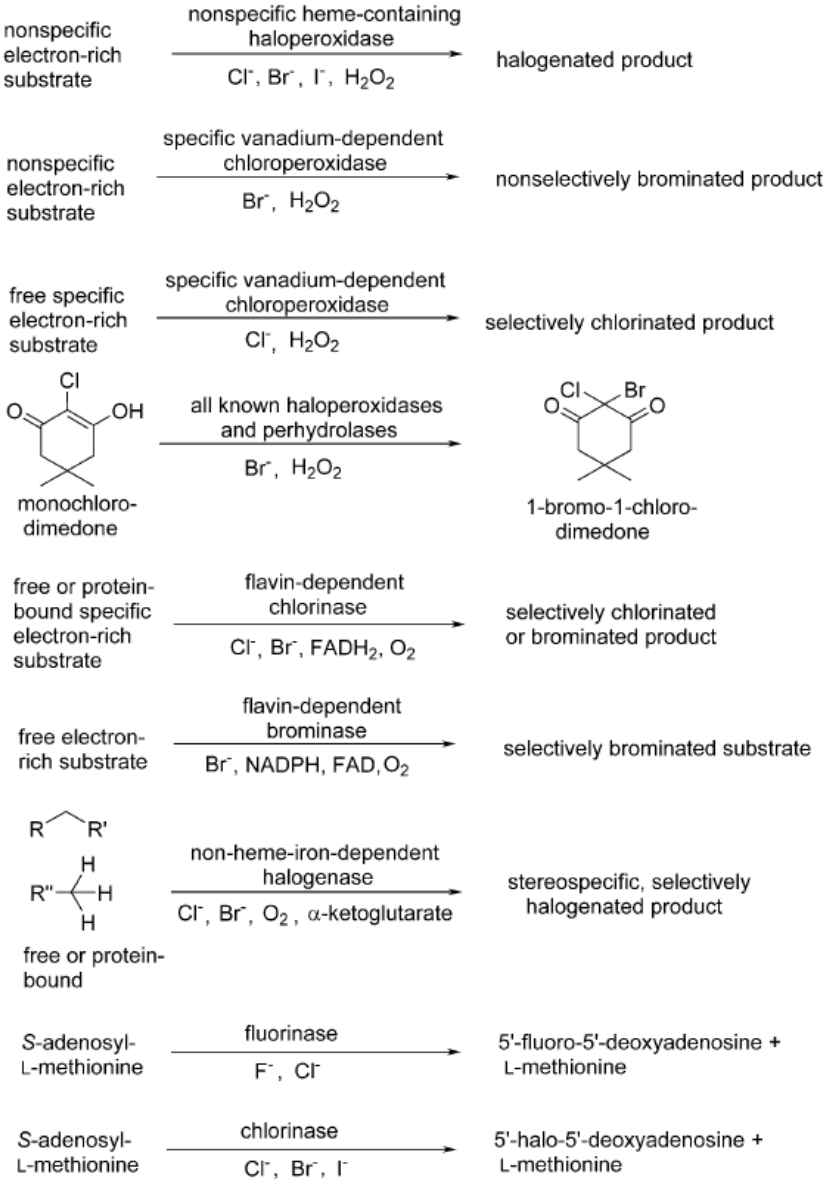


Figure. 1.4 Different types of halogenating enzymes and their reactions. From Weichold *et al*. (2016).

The general mechanism of haloperoxidases involves the catalysis of electrophilic halogenation reactions, whereas the non-heme iron halogenases catalyse radical halogenation reactions and SAM-dependent halogenases catalyse nucleophilic halogenation reactions (Butler and Sandy, 2009). Haloperoxidases have traditionally been named after the most electrophilic/electronegative halide they can readily oxidise (Ohshiro *et al.*, 2004). Therefore, iodoperoxidases (IPO) oxidize only iodide by H_2O_2 whereas CPOs oxidize chloride, bromide and iodide by H_2O_2. The overall stoichiometry of the haloperoxidase reaction is one halogenated product produced at the expense of one equivalent of H_2O_2 (Fig. 1.5 Reaction 1) (Butler and Sandy, 2009). Besides that, H_2O_2 lacks the thermodynamic potential to oxidize fluoride, thus explaining inexistence of fluroperoxidase (Butler and Sandy, 2009).

One class of halogenase, the non-heme iron enzymes catalyse radical halogenation reactions at aliphatic carbon sites by using molecular dioxygen (O_2) to activate halide (Fig. 1.5 Reaction 2). The overall stoichiometry is consumption of one equivalent of α-ketoglutaric acid and O_2 per halogenated equivalent produced (Fig. 1.5 Reaction 2) (Butler and Sandy, 2009). Flavin-dependent halogenases oxidise the halide via the activation of O_2 by means of formation of flavin-bound hydroperoxide (Fig. 1.5 Reaction 3). The reduced flavin $FADH_2$ is provided by an additional enzyme, NAD(P)H-dependent-flavin reductase. This redox cofactor $FADH_2$ is utilised by the two-component flavin reductase/halogenase as a diffusible intermediate between the two enzyme components to catalyse oxidation of the substrate (Vaillancourt *et al.*, 2006). However, direct interaction with the reductase component is not a prerequisite for catalysis when it has been shown that a rhodium organometallic complex can be used for reduced FAD regeneration with formate as the electron donor (Unversucht *et al.*, 2005). SAM-dependent enzymes catalyse nucleophilic substitution, generating 5´-halo- 5´-deoxyadenosine and l-methionine (Fig. 1.5 Reaction 4) (Butler and Sandy, 2009).

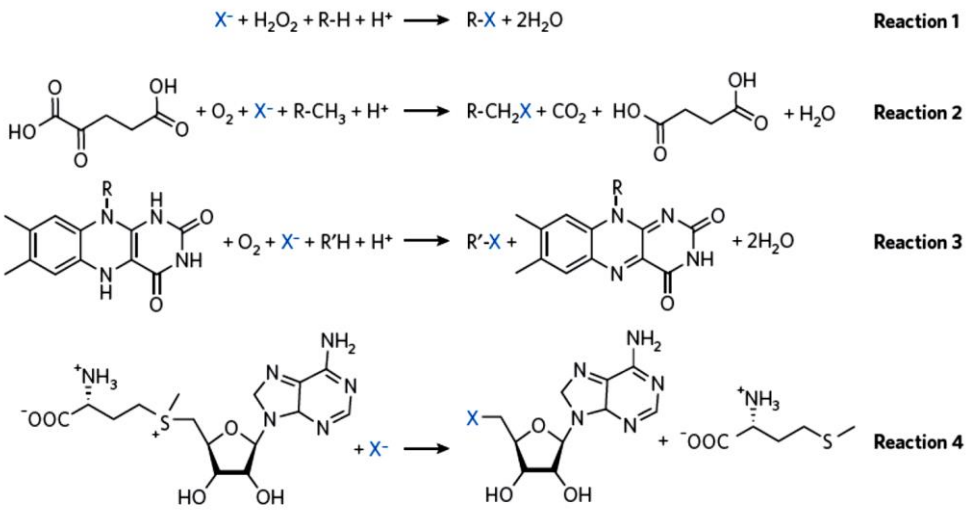


Figure 1.5 Biologically relevant halogenation reaction. **Reaction 1**, which encompasses electrophilic halogenation, is catalysed by haloperoxidases such as the heme-containing or vanadate-containing enzymes. **Reaction 2** is carried out by α-ketoglutarate-dependent nonheme iron halogenases and is characteristic of radical halogenation at unactivated C–H centres. **Reaction 3** is catalysed by flavin-dependent halogenases. **Reaction 4** shows the SAM-dependent halogenase reaction. X, halogen atom; R, R´, alkyl group. From Butler and Sandy (2009).

1.1.3 Tryptophan halogenating enzymes

Many naturally occurring halogenation reactions are catalysed by flavoezymes, the flavin ($FADH_2$)-dependent halogenases (Van Berkel *et al.*, 2006). In particular, halogenated derivatives containing an indole moiety are produced by tryptophan halogenases which are flavin dependent, and various substrates are accepted by these enzymes (Fig. 1.6). The first discovery of flavin-dependent halogenases was reported in the investigation of pyrrolnitrin and chlorotetracyline biosynthesis (Wagner *et al.*, 2009).

In *Pseudomonas fluorescens* BL915, gene disruption experiments have shown that each of the four enzymes (PrnA, B, C and D) complemented the pyrrolnitrin-negative pheynotype of the corresponding deletion mutant (Hammer *et al.*, 1997). Regioselective chlorination of tryptophan in the 7-position of the indole ring is catalysed by PrnA whereas, also with high substrate specificity, PrnC catalyses the chlorination of the phenyl pyrrole derivative monodechloroaminopyrrolnitrin (MCAP) in the 3-position of the pyrrole ring, leading to the formation of aminopyrrolnitrin (Van Pée, 2012). Another Trp-7-halogenase, RebH has been characterised from *Lechevalieria aerocolonigenes* ATCC39243 within the rebeccamycin biosynthetic gene cluster (Wagner *et al.*, 2009). With the notable regioselectivity, RebH catalyses identical conversion of free tryptophan to 7-chlorotryptophan in the biosynthesis of the indolocarbazole natural product rebeccamycin (Yeh *et al.*, 2005). Nonetheless, PrnA recruits a general reductase from cellular metabolism to reduce oxidised cofactor whereas RebH employs a specific reductase RebF derived from the biosynthetic gene cluster (Vaillancourt *et al.*, 2006; Neumann *et al.*, 2008).

Halogenases capable of chlorinating tryptophan at other positions were also found. PyrH, a tryptophan-5-halogenase has been described in pyrroindomycin B biosynthesis of *Streptomyces rugosporus* LL-42D005 (Zehner *et al.*, 2005) whereas the tryptophan-6-halogenase, Thal, was shown to be involved in the production of a plant growth regulating compound thienodolin by *Streptomyces albogriseolus* (van Pée and Zehner, 2003), both thought to catalyse initial steps in the biosynthesis.

A classification of $FADH_2$-dependent haloganases into two subgroups has been proposed based on their substrate specificity, with one group catalysing the halogenation of tryptophan or indole derivatives and the other group acting on pyrrole or phenol derivatives as substrates (van Pée and Zehner, 2003). From the same organism *Pseudomonas fluorescens* BL915 and within the same halometabolite biosynthesis, PrnA and PrnC cannot substitute for each other and their primary structures and gene sequences are fairly distinct with homology below 25% (van Pée and Zehner, 2003; Van Pée, 2012). However, there are two highly conserved motifs, the suggested $FADH_2$ binding site (GXGXXG) and a second motif containing two tryptophan residues (WXWXIP) (van Pée and Zehner, 2003).

The halogenase genes *RebH* and *PyrH* have been cloned from bacteria and have been shown to be functional in transgenic alkaloid producing *Cathanranthus roseus* plants (Runguphan *et al.*, 2010). This provides some insights into cloning the endogenous halogenase gene or genes that encode enzymes catalysing the production of 4-Cl-IAA in pea. The tryptophan halogenases reported in the literature also serve as a pool of candidate genes and blasting them in Genbank, to the genome of *Medicago truncatula*, a close relative of pea as well as Cool Season Food Legume Genome Database could provide a starting point for primer design and subsequent full characterisation. The structure and mechanism of bacterial and fungal tryptophan halogenases are studied and reviewed, thereby shedding some light on plant halogenase with working purification methods and functional bioassay.

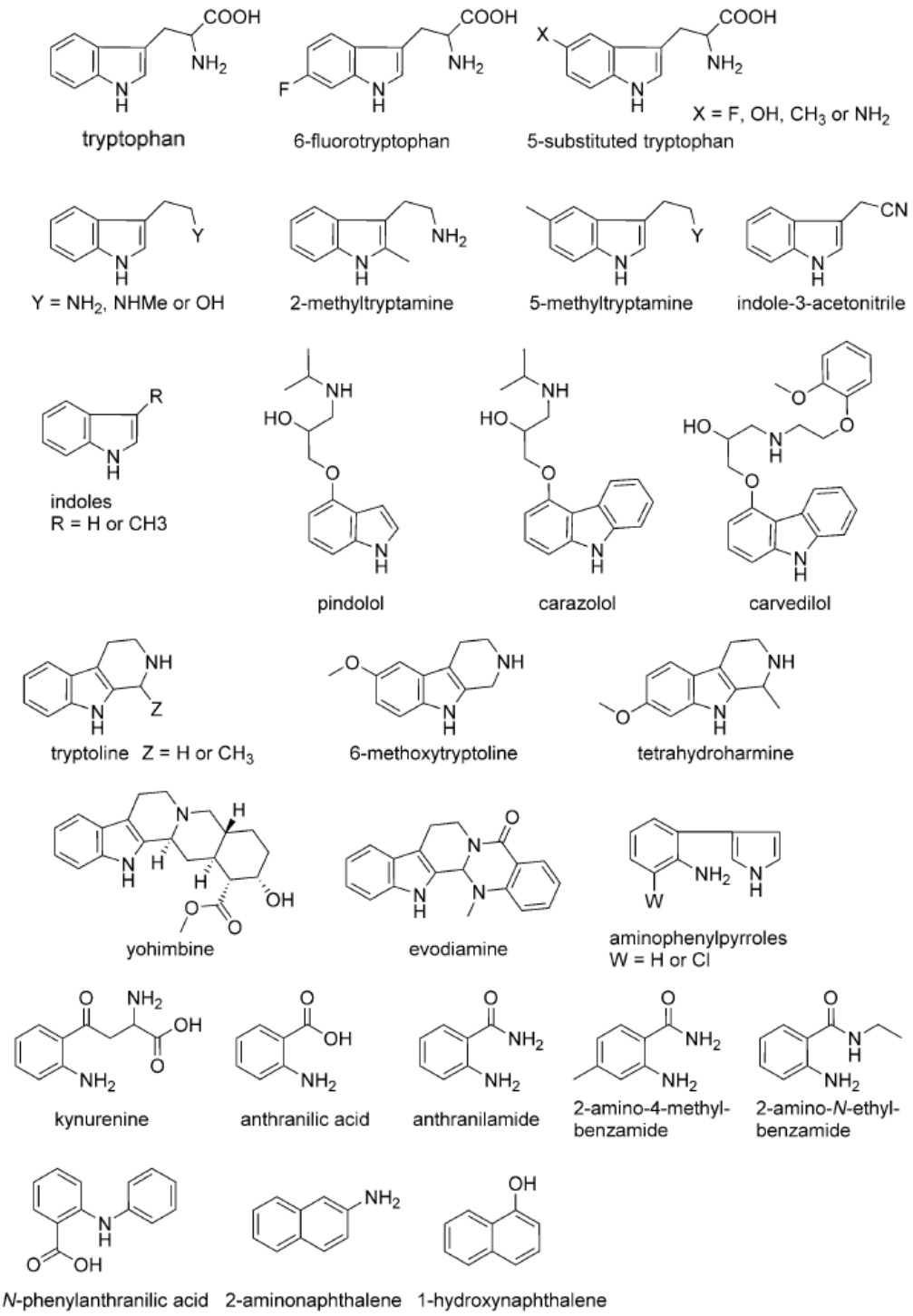


Figure 1.6 Different types of substrates accepted by flavin-dependent tryptophan halogenases. From Weichold *et al*. (2016).

1.2 Aims for this study

Despite the sophisticated characterisation of halogenases in bacteria, fungi and marine algae, little was known about halogenating enzymes in plants. The highly potent halogenated plant hormone, 4-Cl-IAA, was thought to be restricted to the leguminous tribe *Fabeae* (formerly known as tribe *Vicieae*) and intriguingly reported in Scots pine which is not closely related to the Fabaceae. The first aim for this study, as documented in Chapter 2, was to investigate the distribution of 4-Cl-IAA by an extensive survey of the chlorinated hormone and its chlorinated precursor, 4-Cl-tryptophan in the Fabaceae, and by revisiting the questionable occurrence of 4-Cl-IAA reported in Scots pine and studying several other *Pinus* species. Secondly, the study of distribution of 4-Cl-IAA led to the elucidation of evolutionary origin of chlorinated auxin in the Fabaceae, as summarised in Chapter 3. Chapter 4 reports the discovery of non-chlorinating species in the genus *Trigonella* and interesting insights on phylogeny in the Fabaceae. At the onset of this study, genetic sequences for putative pea vanadium-dependent haloperoxidase genes were made available in the pea RNA-seq gene atlas. Therefore, Chapter 5 describes the mutant-based study of a putative vanadium-dependent haloperoxidase in pea. Lastly, Chapter 6 provides a general discussion on the findings of this study, and concluding remark.

Chapter 2 Distribution of Chlorinated Auxins in Plants

2.1 Introduction

The chlorinated form of auxin, 4-chloroindole-3-acetic acid (4-Cl-IAA), is a highly active hormone that is thought to play a key role in early pericarp growth (Reinecke *et al.*, 1995; Reinecke *et al.*, 1999; Ozga *et al.*, 2009). Exogenous 4-Cl-IAA, for example, has been shown to promote the pericarp elongation of de-seeded pea pods (Reinecke *et al.*, 1999). Johnstone *et al.* (2005) reported that 4-Cl-IAA and bioactive gibberellin (GA_3 or GA_1) act synergistically on pericarp growth when applied simultaneously, and a growth regulatory role has been proposed for 4-Cl-IAA through induction of GA biosynthesis and inhibition of ethylene action. In other species, e.g. tomato, the non-chlorinated form of auxin, indole-3-acetic acid (IAA) also stimulates fruit growth via GAs (Serrani *et al.*, 2008; Tang *et al.*, 2015). The chlorinated auxin is mainly found in reproductive structures (Katayama *et al.*, 1988), in which its levels often exceed those of the more widespread IAA (Tivendale *et al.*, 2012). The chlorinated form is thought to be restricted to members of the leguminous tribe *Fabeae* (Reinecke *et al.*, 1999), which includes the genera *Vicia*, *Pisum*, *Lathyrus*, *Lens* and *Vavilovia* (Schaefer *et al.*, 2012). However, there is a curious exception: 4-Cl-IAA has also been reported from both immature and mature seeds of Scots pine, *Pinus sylvestris* (Ernstsen and Sandberg, 1986).

The Fabaceae or Leguminosae, by comprising approximately 20,000 species, is the third largest family of higher plants and their agro-societal importance is evidenced by the area harvested and total production, which are second only to cereal crops (Graminiae or Poaceae) (Gepts *et al.*, 2005). Grain legumes not only provide dietary protein, nitrogen and processed vegetable oil for human consumption but also, together with pasture legumes and woody legumes, fix nitrogen gas from the atmosphere to organic N fertilisers through their hallmark trait of forming symbioses with nitrogen-fixing rhizobial bacteria (Graham and Vance, 2003). This biological nitrogen fixation by legumes may be well fit as the next most fundamentally important process on the planet after photosynthesis (Howieson *et al.*, 2008). In modern farming systems, pasture legumes are the keys to greater animal production

and better crop yields when grown in rotation (Nichols *et al.*, 2012). Apart from the widely grown traditional European perennial pasture legumes, white clover (*Trifolium repens*) and lucerne (*Medicago sativa*), annual subterranean clover (*T. subterraneum*) and annual medics (*Medicago* species) have been developed by plant breeders over the years. To date, farmers can readily access a suite of annual and perennial pasture legume species as well as various cultivars within particular species best adapted to different environmental conditions.

Year 2016 has been declared by the 68th United Nation General Assembly the International Year of Pulses and such implementation has been facilitated by The Food and Agriculture Organization of the United Nations (FAO). This highlights the increasing demand and importance of legumes, together with cereal and other food crop cultivations, to feed the 9 billion human population by 2050 with a 70% rise of global food production. Apart from availability of new and improved cultivars, readily access to scientific datasets is essential to both farmers and scientists. For instance, valuable information such as results of field experiment of grain legumes conducted all over the world have been made available as a database via Dryad Digital Repository (Cernay *et al.*, 2016).

In this study, the distribution of 4-Cl-IAA and 4-Cl-Trp in the Fabaceae was investigated by monitoring these compounds in the seeds of representative species spanning the phylogeny of the family. Most of these species have not been previously tested for the presence of the chlorinated compounds. In addition, the reported occurrence of 4-Cl-IAA outside the Fabaceae, namely in the pine, *Pinus sylvestris* and in several other *Pinus* species were re-examined. The endogenous levels of 4-Cl-IAA in both vegetative tissues and seeds of *Vicia faba* were also examined to address the question of whether or not 4-Cl-IAA is largely restricted to seeds (Pless *et al.*, 1984).

2.2 Materials and methods

2.2.1 Plant materials

Medicago truncatula cv. Jemalong, *Vicia faba*, *Vicia tetrasperma*, *Pisum elatius var. elatius* and *Cicer arietinum* were grown in glasshouse conditions as described previously (Jager *et al.*, 2007). Young seeds or fresh seed pods of other legume species, including *Trifolium repens*, *T. micranthum*, *T. subterraneum*, *Melilotus indicus*, *Indigofera australis*, *Clianthus puniceus*, *Glycine clandestine*, *Hardenbergia comptoniana* and *Pultenaea juniperina* were collected from plants grown in the field (Hobart and Kingston, Tasmania). Dry seeds of *Medicago truncatula* cv. Jemalong, *Lens ervoides*, *Lens nigricans*, *Lotus japonicus*, *Arachis hypogaea*, *Lupinus angustifolius* and *Pinus sylvestris* were obtained from commercial sources and some seeds from the obtained seedlot were grown in glasshouse conditions to confirm their viability. Mature seeds of *Pinus radiata* were harvested from cones collected from wild plants (Kingston, Tasmania). Immature seeds were extracted from young cones of *Pinus flexilis*, *Pinus pinea*, *Pinus parviflora* 'Glauca' and *Pinus parviflora* 'Shikoku-goyo' grown in the Royal Tasmanian Botanical Gardens, Hobart, Tasmania.

2.2.2 Extract preparation for the detection and quantification of compounds

For the extraction and quantification of IAA and 4-Cl-IAA from young, fresh tissues, 0.3-2.5g of tissue was weighed (±0.0001 g FW), placed into a falcon tube with four volumes of cold (-20°C) extraction solvent (80% methanol in water (v/v), with BHT (250 mg/L). The tissue was then homogenised and held at 4°C overnight to extract. For each species, the supernatant was then divided in half, in order to conduct two separate analyses: one for detection of the compounds of interest, and one for quantification of these compounds. For the former, no labelled internal standards were added; and for the latter, [13C6] IAA (Cambridge Isotope Laboratories) and [D4] 4-Cl-IAA (supplied by Prof. Jerry Cohen, Department of Horticultural Science, University of Minnesota) were added as internal standard. The samples were reduced under vacuum at 35°C, taken up in 2% acetic acid in water (v/v), and partitioned twice against diethyl ether. After drying the ether, samples were taken up in 1% acetic acid in water (v/v) and centrifuged for 5 min at 13 000 g. Aliquots were then taken for analysis by UPLC-MS/MS as described below previously (Tivendale *et al.*,

2012). The [D_4]4-Cl-IAA signal detected in this experiment contained a small (6%) contribution from [37Cl, D_2]4-Cl-IAA

In viable, dry seeds, the total levels of each of IAA and 4-Cl-IAA were monitored, including both free acids and conjugated forms. IAA and 4-Cl-IAA were extracted as described by above, but 65% isopropanol in water (v/v), with BHT (250mg/L) was used as the extraction solvent. The supernatant of the extraction is then subjected to a harsh hydrolysis condition using strong base such as sodium hydroxide with incubation at 100℃ to hydrolyse auxin conjugates (ester-linked or amide-linked IAA), therefore allowing the quantification of total auxin content (free + ester-linked + amide-linked IAA or 4-Cl-IAA). From the supernatant, hydrolysis of conjugated IAA and 4-Cl-IAA was carried out using the method described by Symons *et al*. (2002).

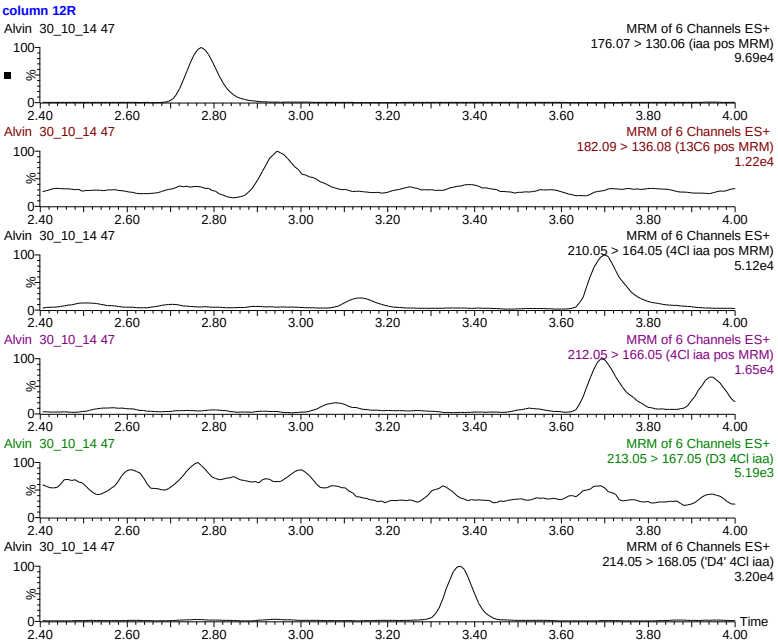
For the extraction of Trp and 4-Cl-Trp, tissue was weighed, homogenised and extracted as described above, but distilled water with BHT (250mg/L) was the usual extraction solvent.

2.3 Results

IAA and 4-Cl-IAA were successfully detected and quantified using the novel detection-quantification methodology based on separate samples of any single tested species adopted in this study (Fig. 2.1). When fresh pods or young seeds are harvested, a mixture of pods or seeds at different developmental stage (i.e. age) is pooled into a single sample extraction. This approach along with highly sensitive UPLC-MS/MS aim to remove the age factor from the levels of hormones. UPLC-MS/MS analyses revealed that IAA was an endogenous compound in all species tested (Table 2.1). 4-Cl-IAA was detected in *T. repens*, *T. micranthum*, *T. subterraneum*, *M. indicus*, *M. truncatula* cv. Jemalong, *Lens ervoides*, *Lens nigricans*, *Vicia tetrasperma*, *Pisum elatius var. elatius* (Fig. 2.2-2.7 and Table 2.1) on the basis of retention times and fragmentation pattern of authentic IAA, [$^{13}C_6$]-IAA, 4-Cl-IAA and [D_4]-4-Cl-IAA (Figs. 2.8 and 2.9). Endogenous 4-Cl-Trp as evidenced by fragmentation pattern shown in Figures 2.10 and 2.11 was detected in *T. repens* (Fig. 2.12), *M. truncatula* (Fig. 2.13), *M. indicus* (Fig. 2.14) but not in *C. arietinum* (Fig. 2.15). Retention times for most species tested range from 1.15 to 1.2 minutes.

Young leaves of *Vicia faba* was shown to contain very low level (3ng.gFW^{-1}) of 4-Cl-IAA (Fig. 2.16) but no detectable level of the same compound in old leaves (Table 2.1). Limited evidence of 4-Cl-IAA was detected in the young leaves of *Trifolium repens* (Fig. 2.17). 4-Cl-IAA was not detected as endogenous compounds by UPLC-MS/MS in Scots pine (*Pinus sylvestris*) and other *Pinus* species (Fig. 2.3 and Table 2.2). Fresh, young pod of *T. repens* was bisected to reveal the developing seeds that are occupying majority of the space within the pod (Fig. 2.18).

A



B

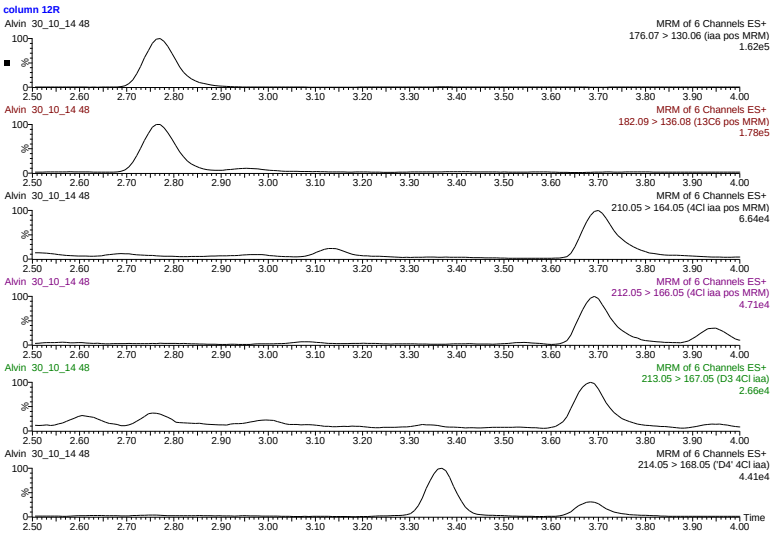


Figure 2.1 UPLC-MS chromatograms (MRM mode) obtained from *Medicago truncatula* (dry seeds, after hydrolysis), showing separate samples subjected to (A) detection and (B) quantification of the compounds of interest without and with the addition of [$^{13}C_6$]-IAA and [D_2]-4-Cl-IAA as internal standards. Peak intensities (normalised) are plotted against retention time; the height of the strongest peak is shown at the right. Transitions: endogenous IAA, 176-130; [$^{13}C_6$]-IAA, 182-136; endogenous 4-Cl-IAA, 210-164; [$^{2}H_2$]4-Cl-IAA, 212-166; [$^{2}H_3$]4-Cl-IAA, 213-167; [$^{2}H_4$]4-Cl-IAA, 214-168.

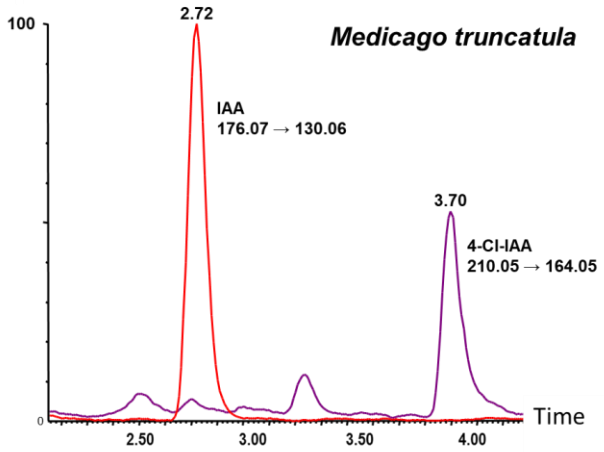


Figure 2.2 UPLC-MS chromatograms (MRM mode) obtained from *Medicago truncatula* (dry seeds, after hydrolysis), showing the presence of endogenous IAA (red channel, MRM transition mass-to-charge ratio, 176.07–130.06; retention time = 2.72 min) and 4-Cl-IAA (purple channel, MRM transition mass-to-charge ratio, 210.05–164.05; retention time = 3.70 min). The products were identified by comparing the retention times and mass spectra with those of authentic IAA and 4-Cl-IAA. Retention times were subjected to minor adjustment to compensate for run-to-run variation.

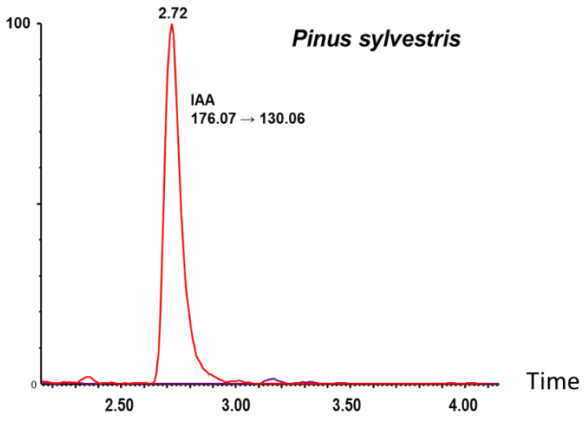


Figure 2.3 UPLC-MS chromatograms (MRM mode) obtained from Scots pine, *Pinus sylvestris* (dry seeds, after hydrolysis), showing the presence of endogenous IAA (red channel, MRM transition mass-to-charge ratio, 176.07–130.06; retention time = 2.72 min) but absence of 4-Cl-IAA (purple channel, MRM transition mass-to-charge ratio, 210.05–164.05; retention time = 3.70 min). The products were identified by comparing the retention times and mass spectra with those of authentic IAA and 4-Cl-IAA. Retention times were subjected to minor adjustment to compensate for run-to-run variation.

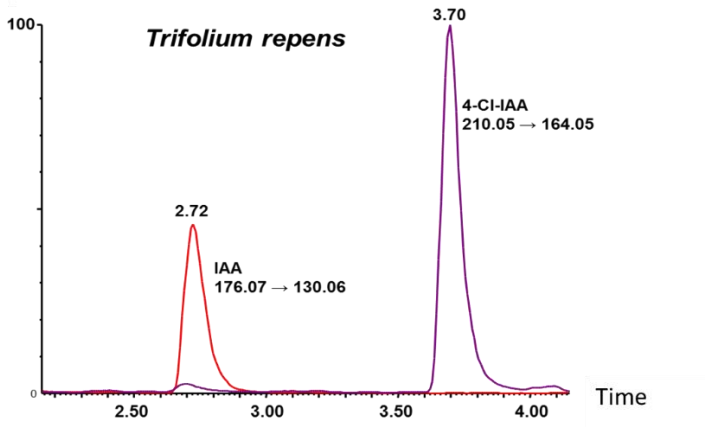


Figure 2.4 UPLC-MS chromatograms (MRM mode) obtained from *Trifolium repens* (fresh pods, containing seeds), showing the presence of endogenous IAA (red channel, MRM transition mass-to-charge ratio, 176.07–130.06; retention time = 2.72 min) and 4-Cl-IAA (purple channel, MRM transition mass-to-charge ratio, 210.05–164.05; retention time = 3.70 min). The products were identified by comparing the retention times and mass spectra with those of authentic IAA and 4-Cl-IAA. Retention times were subjected to minor adjustment to compensate for run-to-run variation.

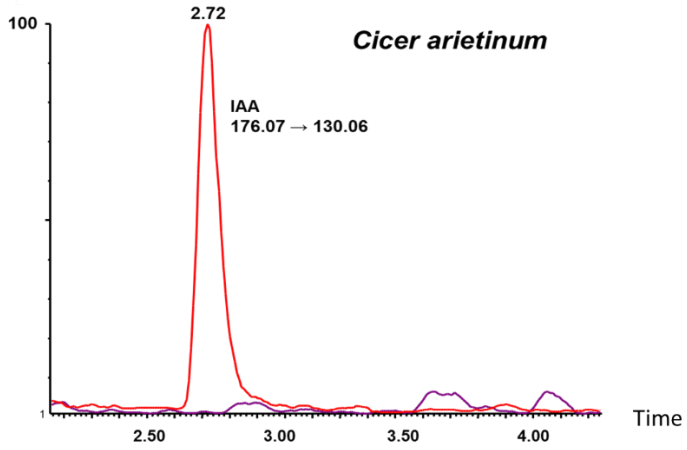


Figure 2.5 UPLC-MS chromatograms (MRM mode) obtained from *Cicer arietinum* (young seeds), showing the presence of endogenous IAA (red channel, MRM transition mass-to-charge ratio, 176.07–130.06; retention time = 2.72 min) but absence of 4-Cl-IAA (purple channel, MRM transition mass-to-charge ratio, 210.05–164.05; retention time = 3.70 min). The products were identified by comparing the retention times and mass spectra with those of authentic IAA and 4-Cl-IAA. Retention times were subjected to minor adjustment to compensate for run-to-run variation.

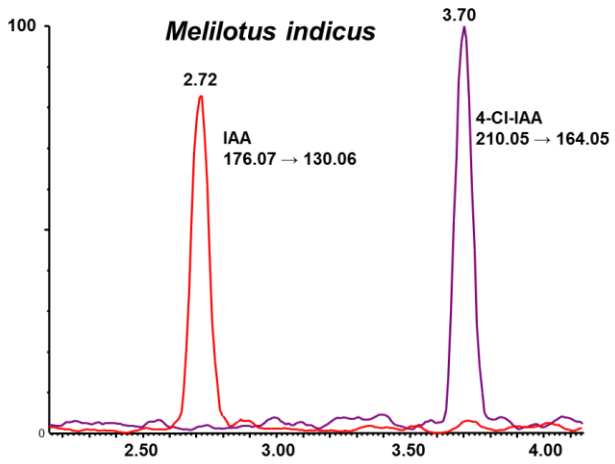


Figure 2.6 UPLC-MS chromatograms (MRM mode) obtained from *Melilotus indicus* (fresh pods, containing seeds), showing the presence of endogenous IAA (red channel, MRM transition mass-to-charge ratio, 176.07–130.06; retention time = 2.72 min) and 4-Cl-IAA (purple channel, MRM transition mass-to-charge ratio, 210.05–164.05; retention time = 3.70 min). The products were identified by comparing the retention times and mass spectra with those of authentic IAA and 4-Cl-IAA. Retention times were subjected to minor adjustment to compensate for run-to-run variation.

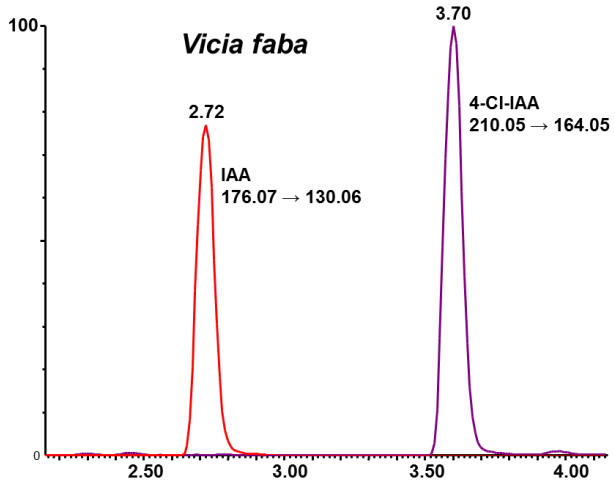


Figure 2.7 UPLC-MS chromatograms (MRM mode) obtained from broad bean, *Vicia faba* (young seeds), showing the presence of endogenous IAA (red channel, MRM transition mass-to-charge ratio, 176.07–130.06; retention time = 2.72 min) and 4-Cl-IAA (purple channel, MRM transition mass-to-charge ratio, 210.05–164.05; retention time = 3.70 min). The products were identified by comparing the retention times and mass spectra with those of authentic IAA and 4-Cl-IAA. Retention times were subjected to minor adjustment to compensate for run-to-run variation.

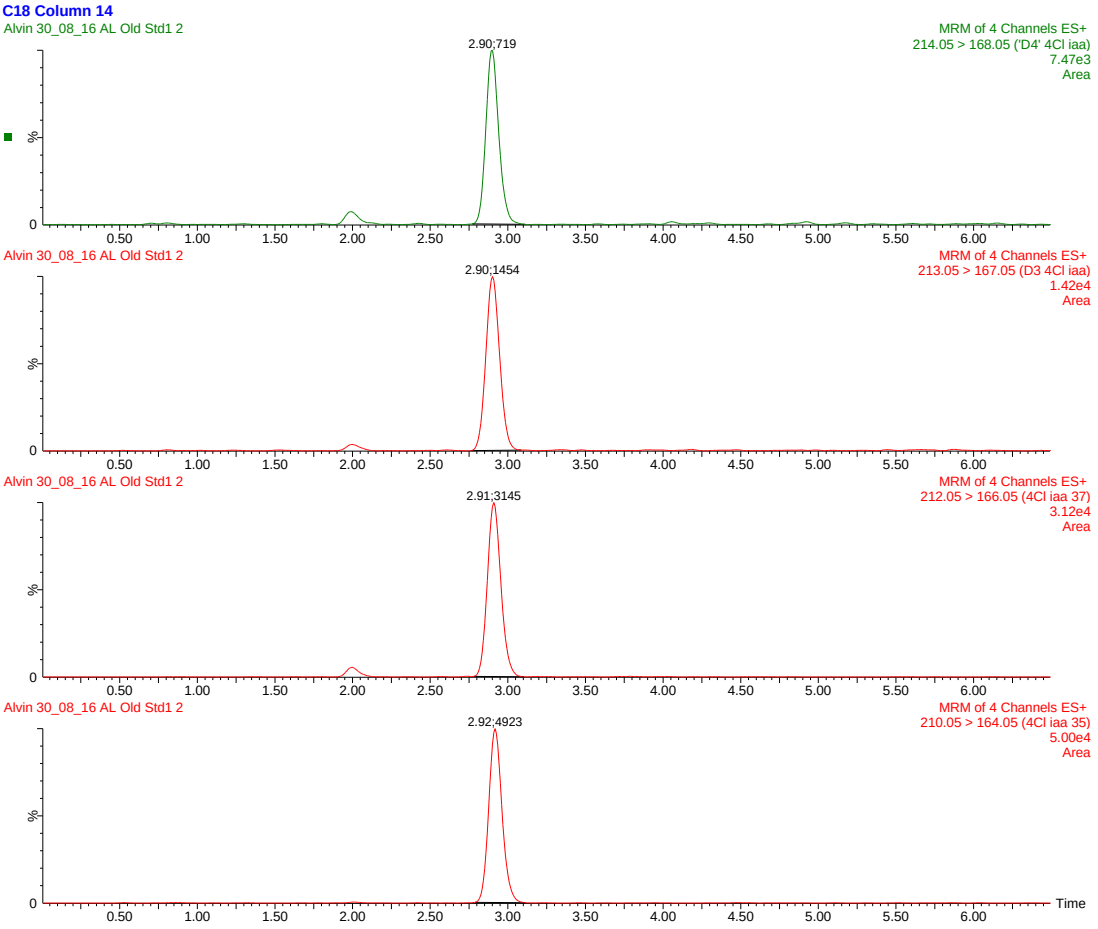


Figure 2.8 UPLC-MS chromatogram (MRM for 4-Cl-IAA and [D_4]-4-Cl-IAA) of authentic 4-Cl-IAA and [D_4]-4-Cl-IAA in methanol. Retention times may be subject to run-to-run variation.

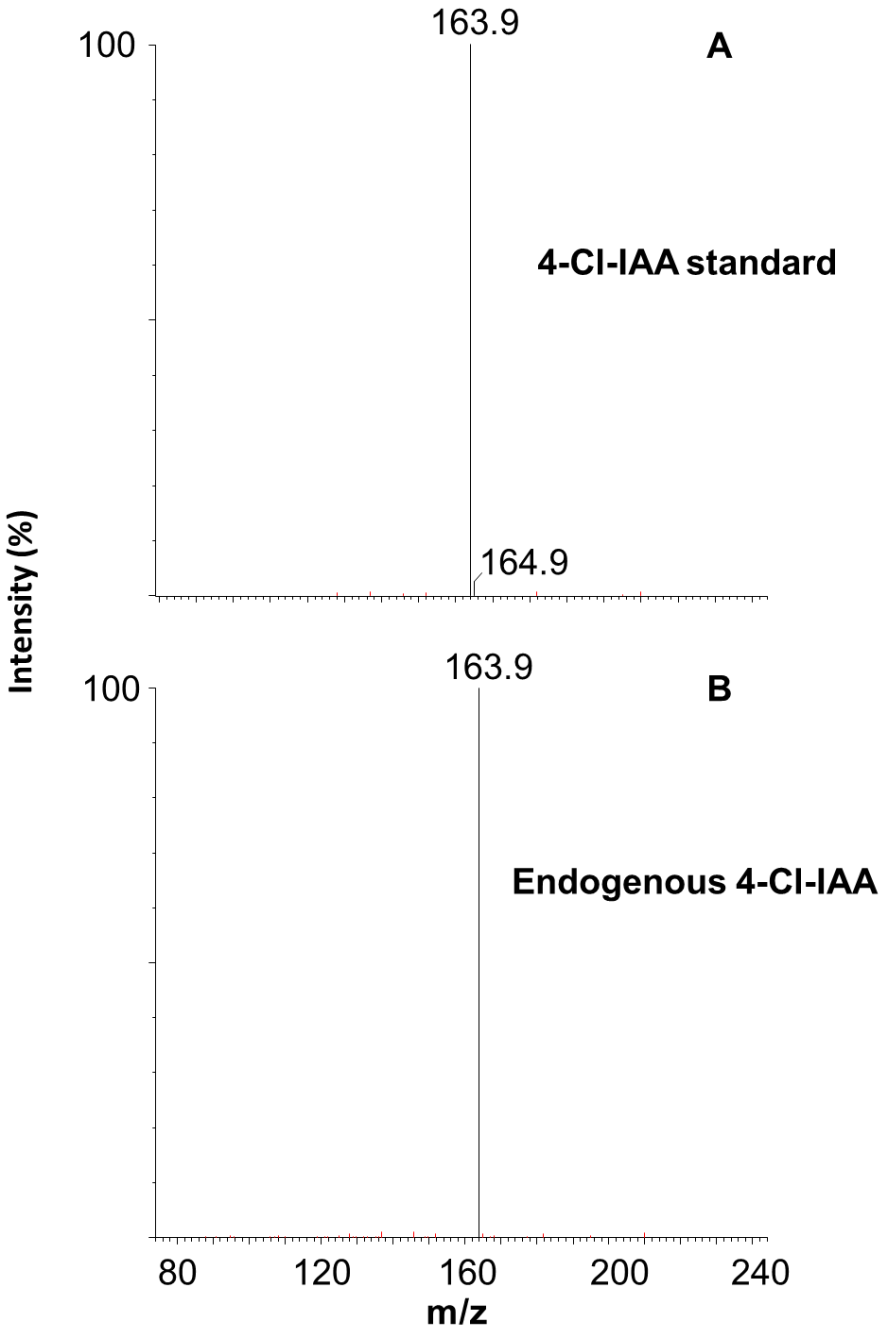


Figure 2.9 Product ion spectrum from ^{35}Cl 4-Cl-IAA. (A) Product ions of M+H (m/z 210) from 4-Cl-IAA standard. (B) Endogenous 4-Cl-IAA from *Trifolium repens*.

Table 2.1 Endogenous levels of 4-Cl-IAA and IAA in 26 representative species from the family Fabaceae. (n.d., not detected)

Species	**Phylogenetic clade**	**Tissue type**	**IAA ($ng.g^{-1}$)**	**4-Cl-IAA ($ng.g^{-1}$)**
Medicago truncatula	*Trifolieae*	Dry seeds	26*	85*
Trifolium repens	*Trifolieae*	Fresh pods, containing seeds	36	1388
		Young leaves	3	n.d.
Trifolium micranthum	*Trifolieae*	Fresh pods, containing seeds	47	569
Trifolium subterraneum	*Trifolieae*	Fresh pods, containing seeds	100	243
Melilotus indicus	*Trifolieae*	Fresh pods, containing seeds	7	39
Vicia faba	*Fabeae*	Young seeds	175	885
		Old seeds	73	91
		Young leaves	16	3
		Old leaves	9	n.d.
Vicia tetrasperma	*Fabeae*	Young seeds	8	97
Pisum elatius var. elatius	*Fabeae*	Young seeds	79	1873
Lens ervoides	*Fabeae*	Dry seeds	1325	143
Lens nigricans	*Fabeae*	Dry seeds	258	402
Cicer arietinum	*Cicereae*	Young seeds	14	n.d.

Clianthus puniceus	*Galegeae*	Young seeds	343	n.d.
Indigofera australis	Indigoferoid	Fresh pods, containing seeds	11	n.d.
Lotus japonicus	Robinioid	Dry seeds	98*	n.d.*
Glycine clandestine	Millettioid	Young seeds	340	n.d.
Hardenbergia comptoniana	Millettioid	Young seeds	2066	n.d.
Pultenaea juniperina	Mirbelioid	Young seeds	274	n.d.
Arachis hypogaea	Dalbergioid	Dry seeds	137*	n.d.*
Lupinus angustifolius	Gesnistoid	Dry seeds	30*	n.d.*

*Total level of IAA or 4-Cl-IAA, including both free acids and conjugated forms.

Table 2.2 Endogenous levels of 4-Cl-IAA and IAA of 6 species in *Pinus*. (n.d., not detected)

Species	**Tissue Type**	**IAA (ng.g^{-1})**	**4-Cl-IAA (ng.g^{-1})**
Pinus sylvestris	Dry seeds	320*	n.d.*
Pinus flexilis	Immature seeds	74	n.d.
Pinus pinea	Mature seeds	38	n.d.
Pinus radiata	Mature seeds	3-6	n.d.
Pinus parviflora 'Glauca'	Immature seeds	71	n.d.
Pinus parviflora 'Shikoku-goyo'	Immature seeds	121	n.d.

*Total level of IAA or 4-Cl-IAA, including both free acids and conjugated forms.

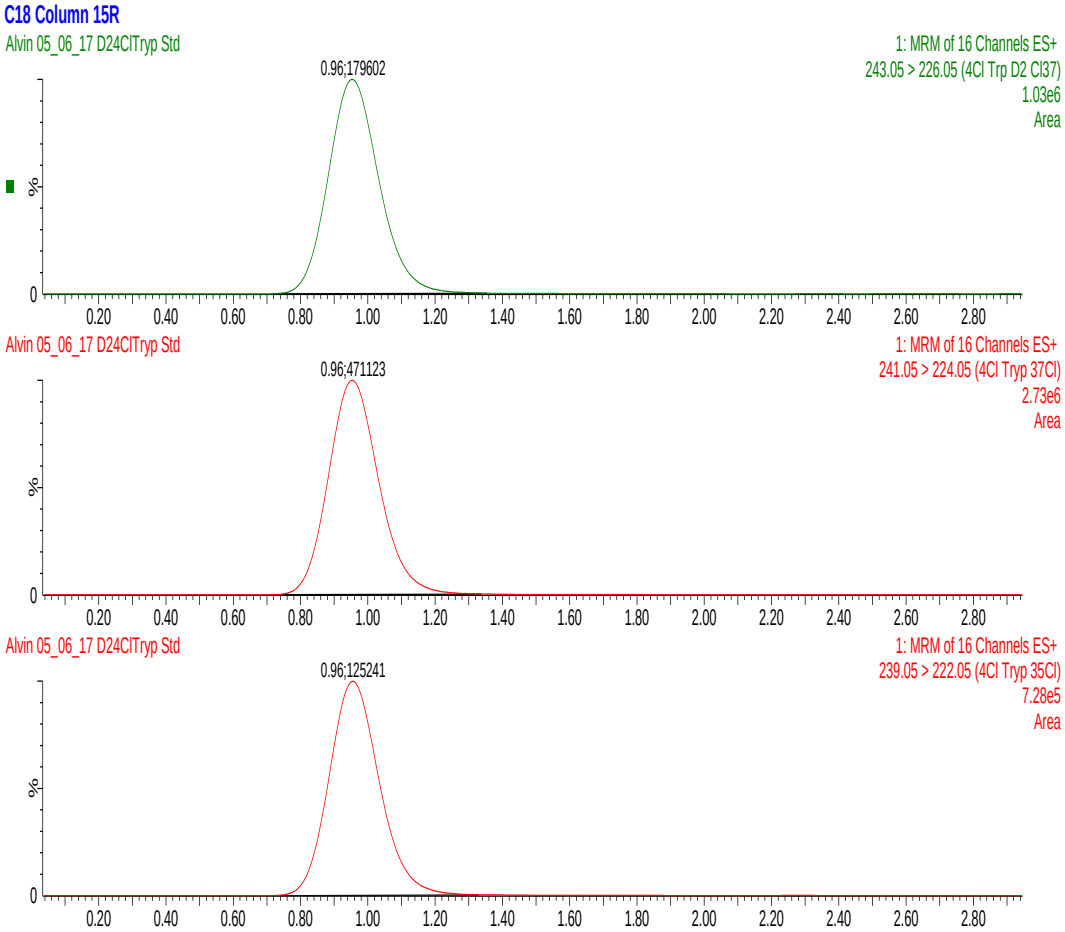


Figure 2.10 UPLC-MS chromatogram (MRM for 4-Cl-Trp and $^{2}H_4$-4-Cl-Trp) of authentic 4-Cl-Trp and D_2-4-Cl-Trp in H_2O. Retention times may be subject to run-to-run variation.

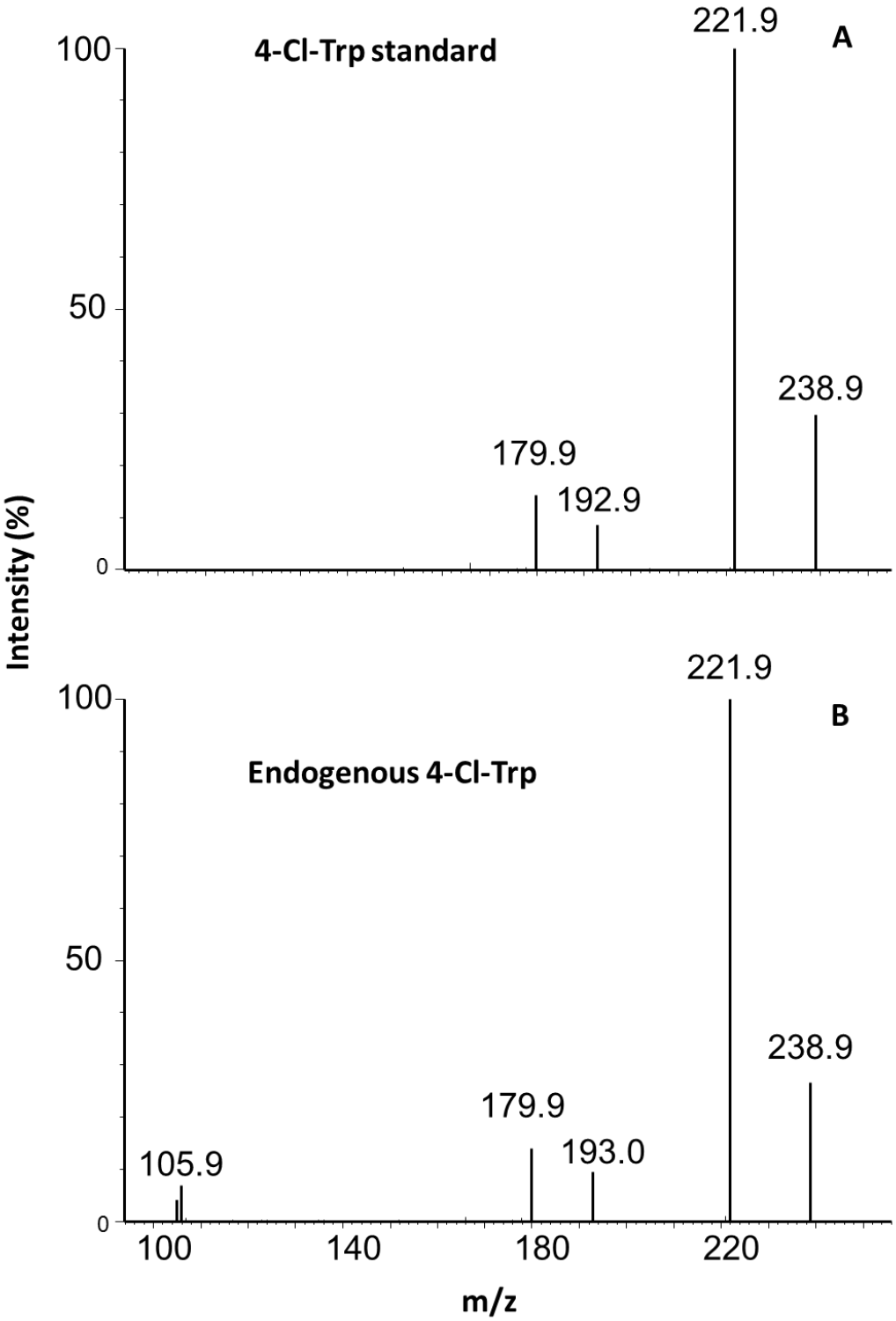


Figure 2.11 Product ion spectrum from ^{35}Cl 4-Cl-Trp. (A) Product ions of M+H (m/z 239) from 4-Cl-Trp standard. (B) Endogenous 4-Cl-Trp from *Trifolium repens*.

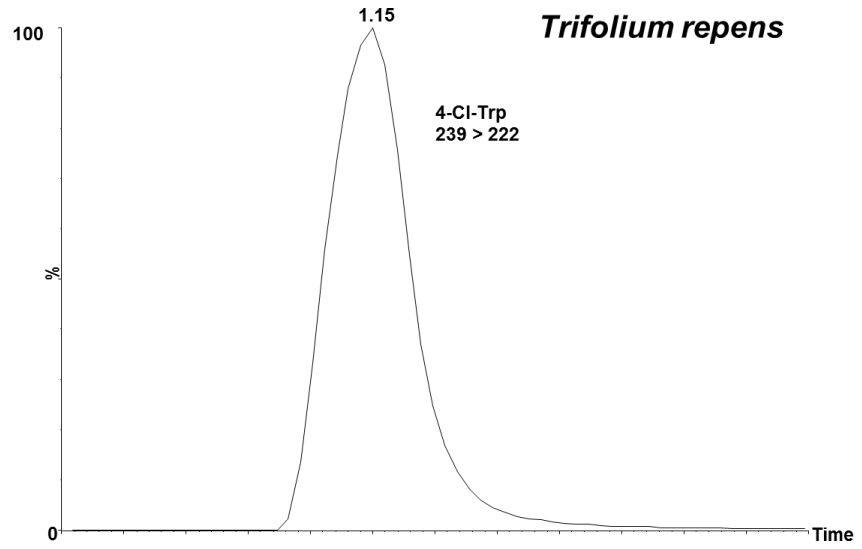


Figure 2.12 UPLC-MS chromatograms (MRM mode) obtained from *Trifolium repens* (fresh pods, containing seeds), showing the presence of endogenous 4-Cl-Trp (MRM transition mass-to-charge ratio, 239 – 222; retention time = 1.15 min). The products were identified by comparing the retention times and mass spectra with those of authentic 4-Cl-Trp. Retention times were subjected to minor adjustment to compensate for run-to-run variation.

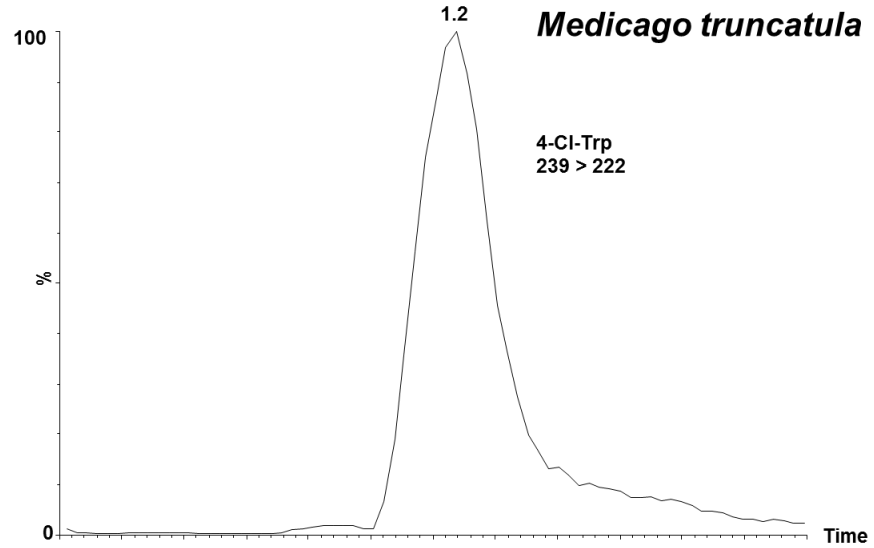


Figure 2.13 UPLC-MS chromatograms (MRM mode) obtained from *Medicago truncatula* (fresh burrs, containing seeds), showing the presence of endogenous 4-Cl-Trp (MRM transition mass-to-charge ratio, 239 – 222; retention time = 1.15 min). The products were identified by comparing the retention times and mass spectra with those of authentic 4-Cl-Trp. Retention times were subjected to minor adjustment to compensate for run-to-run variation.

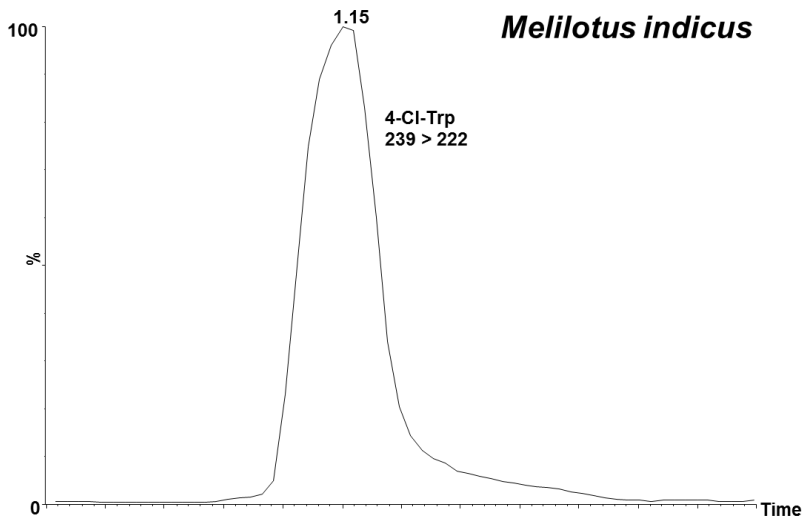


Figure 2.14 UPLC-MS chromatograms (MRM mode) obtained from *Melilotus indicus* (fresh pods, containing seeds), showing the presence of endogenous 4-Cl-Trp (MRM transition mass-to-charge ratio, 239 – 222; retention time = 1.15 min). The products were identified by comparing the retention times and mass spectra with those of authentic 4-Cl-Trp. Retention times were subjected to minor adjustment to compensate for run-to-run variation.

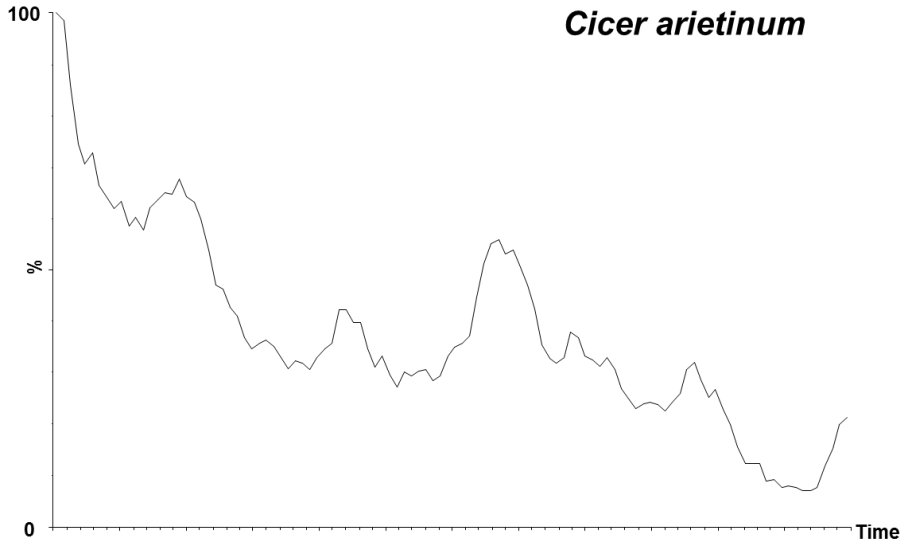


Figure 2.15 UPLC-MS chromatograms (MRM mode) obtained from *Cicer arietinum* (young seeds), showing the absence of endogenous 4-Cl-Trp (MRM transition mass-to-charge ratio, 239 – 222; retention time = 1.15 min). The products were identified by comparing the retention times and mass spectra with those of authentic 4-Cl-Trp. Retention times were subjected to minor adjustment to compensate for run-to-run variation.

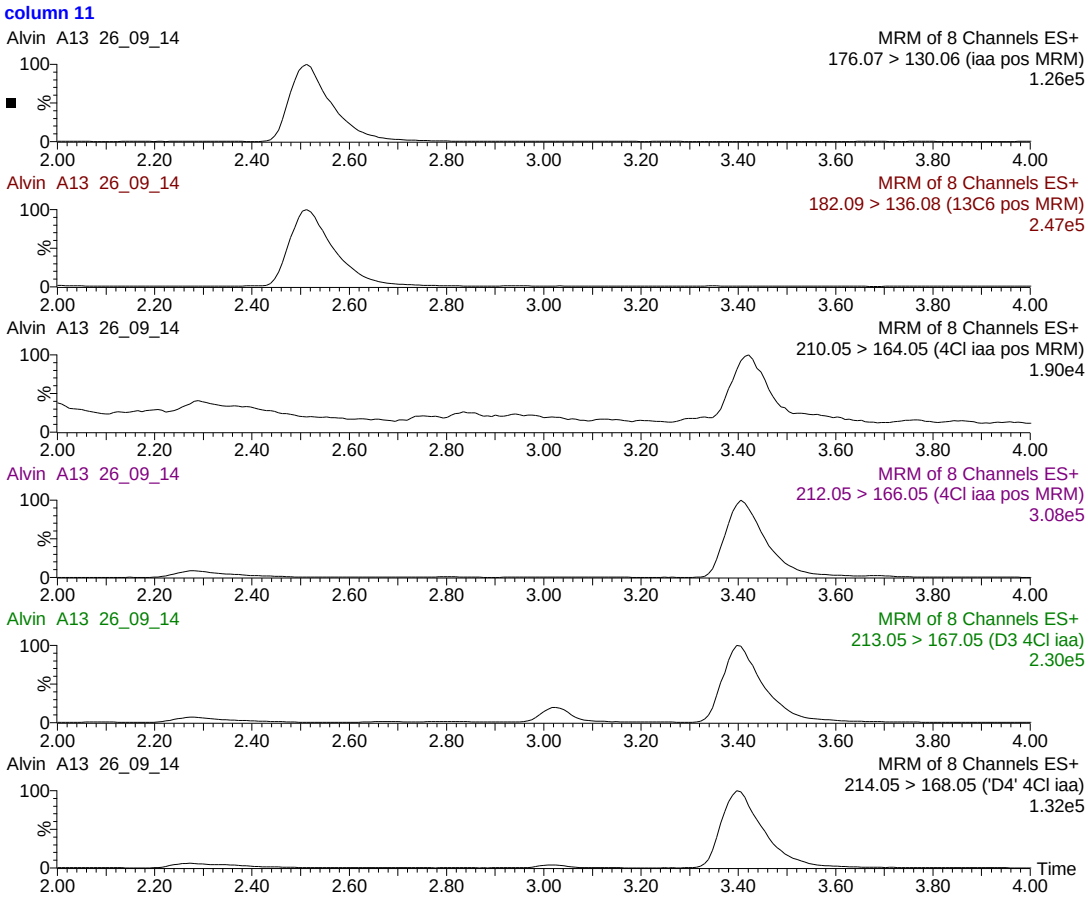


Figure 2.16 UPLC-MS chromatograms (MRM mode) obtained from young leaves of broad bean, *Vicia faba*, showing the presence of endogenous IAA (MRM transition mass-to-charge ratio, 176.07–130.06; retention time = 2.52 min) and 4-Cl-IAA (MRM transition mass-to-charge ratio, 210.05–164.05; retention time = 3.42 min). Peak intensities (normalised) are plotted against retention time; the height of the strongest peak is shown at the right. Transitions: endogenous IAA, 176-130; [$^{13}C_6$]-IAA, 182-136; endogenous 4-Cl-IAA, 210-164; [$^{2}H_2$]4-Cl-IAA, 212-166; [$^{2}H_3$]4-Cl-IAA, 213-167; [$^{2}H_4$]4-Cl-IAA, 214-168.

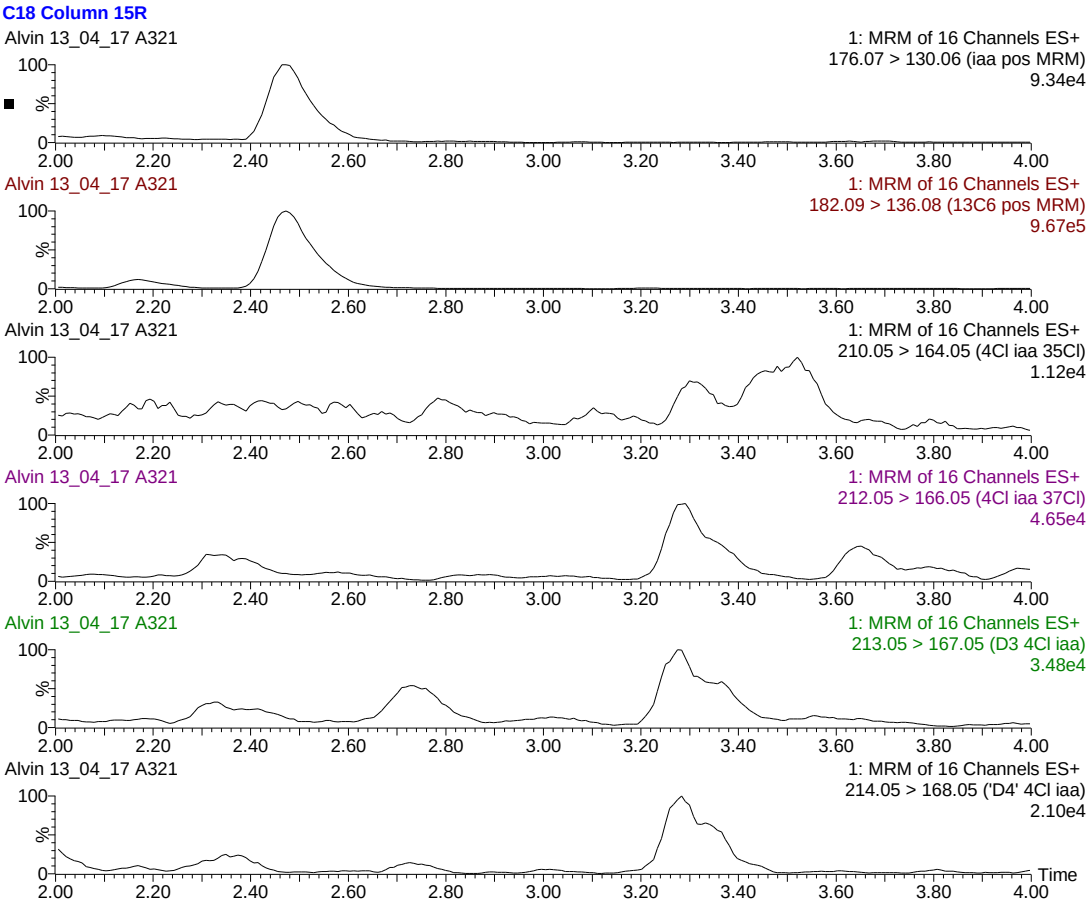


Figure 2.17 UPLC-MS chromatograms (MRM mode) obtained from young leaves of *Trifolium repens*, showing the presence of endogenous IAA (MRM transition mass-to-charge ratio, 176.07–130.06; retention time = 2.52 min) but limited evidence of 4-Cl-IAA (MRM transition mass-to-charge ratio, 210.05–164.05; retention time = 3.30 min). Peak intensities (normalised) are plotted against retention time; the height of the strongest peak is shown at the right. Transitions: endogenous IAA, 176-130; [$^{13}C_6$]-IAA, 182-136; endogenous 4-Cl-IAA, 210-164; [$^{2}H_2$]4-Cl-IAA, 212-166; [$^{2}H_3$]4-Cl-IAA, 213-167; [$^{2}H_4$]4-Cl-IAA, 214-168.

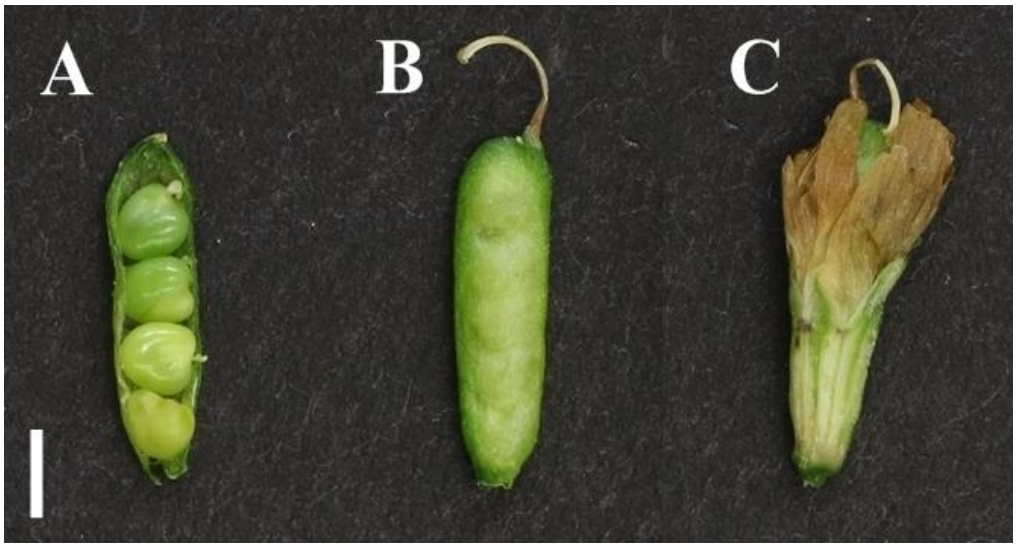


Figure 2.18 Fruit of *Trifolium repens*. (A) Seeds in a bisected seed pod. (B) Whole seed pod. (C) Floret covering a single seed pod. Scale bar, 1mm.

2.4 Discussion

Twenty-six species from the Fabaceae were screened for chlorinated auxin, selected to represent all major clades of the family (Wojciechowski *et al.*, 2004). The range of leguminous species that contain 4-Cl-IAA was extended, by detecting this compound in reproductive structures of *Trifolium repens*, *T. subterraneum*, *T. micranthum*, *Melilotus indicus* and *Medicago truncatula* all from outside the tribe *Fabeae*. Identification was based firstly on the correct LC retention time and fragmentation patterns, as indicated by comparison with standard forms of 4-Cl-IAA (Figs. 2.8 and 2.9). Secondly, peaks for the expected mass transitions, 210 to 164 and 212 to 166, were observed, with the latter comprising approximately 33% of the former (data not shown). This ratio of peak areas is diagnostic for compounds containing one chlorine atom (Gross, 2011). The peak corresponding to the transition for 211 to 165 was only 10% of the 210 to 164 peak, as expected from naturally-occurring isotopes. The product scan for 4-Cl-IAA is very simple and direct but also informative and diagnostic as the product scan of 4-Cl-IAA from the samples does match the standard 4-Cl-IAA product scan without showing any other fragments that are not endogenous in the standard. In some cases, dry seeds were analysed (Table 2.1), and in these cases, the extract was hydrolysed to release 4-Cl-IAA and IAA from their conjugated forms.

UPLC-MS/MS analyses revealed that the novel methodology in this study using separate samples for detection of compounds of interest (without the addition of labelled internal standards) and quantification of these compounds (with the addition of labelled internal standards) was successful and effective. In short, the novelty of this approach (that is, with and without internal standard) is that when chlorinated IAA is absent in the sample, we can show by the presence of the internal standard that the absence of endogenous 4-Cl-IAA is not due to loss on work-up. For instance, for detection purpose, when labelled internal standards [$^{13}C_6$]-IAA and [D_2]-4-Cl-IAA were not added, authentic mass spectra of IAA (m/z 176 to 130) and 4-Cl-IAA (^{35}Cl m/z 210 to 164 and ^{37}Cl m/z 212 to 166 in 3:1 diagnostic ratio) were present (Fig. 2.1). For IAA, [$^{13}C_6$]-IAA and endogenous IAA eluted at the same retention time, RT = 2.77 min (Fig. 2.1 B), and therefore, when [$^{13}C_6$]-IAA was not added,

endogenous also eluted at RT = 2.77 min (Fig. 2.1 A). For 4-Cl-IAA, the deuterium atoms caused the deuterated 4-Cl-IAA internal standard to elute slightly earlier than endogenous 4-Cl-IAA (Fig. 2.1 B); RTs of transitions of 210-164 endogenous 4-Cl-IAA > [$^2H_2$]4-Cl-IAA, 212-166 > [$^2H_3$]4-Cl-IAA, 213-167 > [$^2H_4$]4-Cl-IAA, 214-168. This allows much confident confirmation of the presence of a particular compound of interest without the undesired masking effect of unlabelled species from the sometimes, incomplete isotopic labelling of internal standards.

There were very high levels of 4-Cl-IAA in fruits of *T. repens* (which consisted mainly of seeds, Fig. 2.18), and in immature seeds of *P. elatius var. elatius* and *V. faba* as well as in the dry seeds of the four *Ononis* species and two *Lens* species (Table 2.1). In the case of *M. truncatula*, the most convincing evidence for chlorinated auxin was obtained after hydrolysis of dry seeds. Nine species have been shown for the first time here to contain 4-Cl-IAA including *M. truncatula*, *M. indicus*, *T. repens*, *T. subterraneum*, *T. micranthum*, *L. ervoides*, *L. nigricans*, *Vicia tetrasperma* and *Pisum elatius var. elatius*. The level of the chlorinated compound in reproductive structures of these species (except *Lens ervoides*) exceeded that of IAA, as is the case for pea (Tivendale *et al.*, 2012) and *V. faba* (Table 2.1).

Chlorinated auxin was not detected in representative species from the *Galegeae* tribe (*Clianthus puniceus*), the *Robinioids* (*Lotus japonicus*), the *Indigoferoids* (*Indigofera australis*), the *Millettioids* (*Glycine clandestine* and *Hardenbergia comptoniana*), the *Mirbelioids* (*Pultenaea juniperina*), the *Dalbergioids* (*Arachis hypogaea*) or the *Genistoids* (*Lupinus angustifolius*) (Table 2.1). Previously, the 4-Cl-IAA was not detected in a number of *Dalbergioid* species, including *Phaseolus vulgaris* (Hofinger and Böttger, 1979), *Glycine max*, *Vigna catiang* and *Dolichos lablab* (Katayama *et al.*, 1987). Furthermore, a previous finding that 4-Cl-IAA is not detectable in *Cicer arietinum* was confirmed by this study (Fig. 2.15) (Engvild, 1994).

Endogenous 4-Cl-Trp was also detected in *M. truncatula* burrs, as well as the fresh pods containing seeds of *T. repens* and *M. indicus*, providing further evidence provides that Trp or a Trp precursor can be chlorinated (Tivendale *et al.*, 2012; Figs. 2.11-2.13). Again, chlorinated Trp was identified on the basis of retention times and fragmentation pattern, as well as the diagnostic ratio (approximately 3:1) of peak

areas for the molecular ion to M+2 (Gross, 2011) (Figs. 2.10 and 2.11). Without subjected to minor adjustment, retention times may vary slightly between run to run but other diagnostic features in the chromatograms confirm the identity without doubt. Chlorinated Trp was not detected in the Fabaceae species, *C. arietinum* (Fig. 2.15), *C. puniceus*, *I. australis*, *G. clandestine* and *H. comptoniana*, or the *Pinus* species, *P. flexilis* or *P. parviflora*.

Scots pine (*Pinus sylvestris*) is the only species outside the Fabaceae previously reported to contain chlorinated auxin, although the levels reported (only from seeds) were low (Ernstsen and Sandberg, 1986). Due to nature that Scots pine and legumes are not closely related, such claim seems implausible and has prompted us to re-test it ourselves. Apart from the unrelatedness of legumes (angiosperm) and pines (gymnosperm), previous report adopted GC-MS and 7-Cl-IAA as internal standard as 4-Cl-IAA standard was not available in that period of time. In this investigation, UPLC-MS/MS failed to detect any trace of 4-Cl-IAA in dry seeds of Scots pine (Table 2.2; Fig. 2.3). Furthermore, 4-Cl-IAA was not detected in young seeds of a diversity of species from the genus *Pinus*, including *P. flexilis*, *P. parviflora* 'Glauca' and *P. parviflora* 'Shikoku-goyo' nor in mature seeds of *P. pinea* or *P. radiata* (Table 2.2). These results cast serious doubt over the reported presence of 4-Cl-IAA in *Pinus sylvestris* seeds.

In another intriguing report, very high levels (16,000 ng.g(FW)$^{-1}$) of 4-Cl-IAA were documented for young but fully developed leaves of field-grown broad bean plants, *V. faba* (Pless *et al.*, 1984). This is inconsistent with the hypothesis that 4-Cl-IAA is largely restricted to seeds (Katayama *et al.*, 1987). Moreover, the figure reported by Pless *et al.* (1984) exceeds their reported seed content (up to 15,000 ng.g(FW)$^{-1}$). To investigate these findings, the distribution of 4-Cl-IAA in *Vicia faba* and *Trifolium repens* were determined. For *V. faba*, only very low levels were detected in young leaves (approximately 3 ng.g(FW)$^{-1}$) but none in mature leaves. As for *T. repens*, no detectable level of 4-Cl-IAA but only 3 ng.g(FW)$^{-1}$ IAA was found in the young leaves. This low level of chlorinated auxin in young leaves may have derived from the auxin pool in the cotyledon leaves. The levels of 4-Cl-IAA in young and mature

seeds of *V. faba* were quantified as 886 ng.g(FW)$^{-1}$ and 90 ng.g(FW)$^{-1}$ respectively. These data do not support the reported levels of Pless *et al.* (1984), and confirm instead previous evidence (Katayama *et al*. 1988) that seeds contain much higher levels of chlorinated auxin than vegetative tissues.

In conclusion, the evidence presented here extends the range of species that produce chlorinated auxin beyond the *Fabeae*, but restricts it to the Fabaceae. Unequivocal evidence for 4-Cl-IAA and its precursor, 4-Cl-Trp, was obtained from *Medicago truncatula*, and from genera between *Medicago* and the *Fabeae* in phylogenetic terms. However, the finding that chlorinated auxin occurs in *Pinus sylvestris* was unable to be repeated but another four *Pinus* species were further shown to be negative for chlorinated auxin. This exciting discovery of 4-Cl-IAA and 4-Cl-Trp in *Trifolium*, a genus comprised of agriculturally important pasture legumes, and in *Medicago truncatula*, a model species for genetic studies, will potentially be useful to elucidate the biohalogenation in plants. From the phylogenetic point of view, the results also look promising to provide a plausible explanation for the rather enigmatic relationships among taxa within the Fabaceae in the following chapter of the thesis.

Chapter 3 The Single Evolutionary Origin of Chlorinated Auxin in the Fabaceae

3.1 Introduction

Chlorinated auxin is one of the most potent naturally-occurring hormones as demonstrated in various auxin bioassays (Reinecke, 1999). The phylogenetic distribution and evolutionary origins of this hormone, however, have received little attention. The discovery of 4-Cl-IAA in the seeds of *Medicago truncatula*, *Melilotus indicus* and three *Trifolium* species as summarised in Chapter 3 has pushed the number of recognised 4-Cl-IAA-producing leguminous species to 18 and extended the detection of this hormone beyond species of the tribe Fabeae, to which this hormone was previously thought to be restricted (Lam *et al.*, 2015). This study has also demonstrated that Scots pine as well as five other *Pinus* species do not produce chlorinated auxin in their seeds. This finding is regarded as the evidence that chlorinated auxin is restricted to the Fabaceae but not to the Fabeae.

It is surprising that the ability of these species to produce chlorinated auxin had remained unrecognised, given that clovers (*Trifolium* species) have a long cultivation history due to their agricultural significance, especially in temperate regions as forage crops (Smýkal *et al.*, 2015); and that *Medicago truncatula* is a model species with a sequenced genome (Young and Udvardi, 2009). While the phylogenetic relationships of taxa within the Fabaceae have been somewhat enigmatic, the reconstruction of a phylogeny based on the analysis of the plastid MATK gene resolves many well-supported subclades within the legume family (Wojciechowski *et al.*, 2004). A mandatory screening should be performed on representative species from a number of closely related genera such as *Ononis*, *Galega*, *Parochetus* and *Astragalus* for 4-Cl-IAA and –Trp. Also, it is sensible to test an additional species of *Cicer* for endogenous chlorinated compounds as an assurance that the members of the genus are non-chlorinating. The aim of this series of screening experiments is to utilise this greater resolution and clade support to investigate the evolutionary origin of 4-Cl-IAA in the Fabaceae.

3.2 Materials and methods

3.2.1 Plant materials

Dry seeds of *Ononis natrix* and *O. spinosa* were obtained from Australian Pasture Genebank whereas *O. fruticosa*, *O. repens*, *Galega officinalis*, *Parochetus communis*, *Astragalus propinquus* and *A. sinicus* were obtained from commercial sources. Dry seeds of *Cicer echinospermum* were sourced from Agricultural Research Service, United States Department of Agriculture.

3.2.2 Extract preparation for the detection and quantification of compounds

For the extraction and quantification of IAA and 4-Cl-IAA from young, fresh tissues, 0.3-2.5g of tissue was weighed (±0.0001 g FW), placed into a falcon tube with four volumes of cold (-20°C) extraction solvent (80% methanol in water (v/v), with BHT (250 mg/L). The tissue was then homogenised and held at 4°C overnight to extract. For each species, the supernatant was then divided in half, in order to conduct two separate analyses: one for detection of the compounds of interest, and one for quantification of these compounds. For the former, no labelled internal standards were added; and for the latter, [13C6] IAA (Cambridge Isotope Laboratories) and [D4] 4-Cl-IAA (supplied by Prof. Jerry Cohen, Department of Horticultural Science, University of Minnesota) were added as internal standard. The samples were reduced under vacuum at 35°C, taken up in 2% acetic acid in water (v/v), and partitioned twice against diethyl ether. After drying the ether, samples were taken up in 1% acetic acid in water (v/v) and centrifuged for 5 min at 13 000 g. Aliquots were then taken for analysis by UPLC-MS/MS as described below previously (Tivendale *et al*., 2012).

In viable, dry seeds, the total levels of each of IAA and 4-Cl-IAA were monitored, including both free acids and conjugated forms. IAA and 4-Cl-IAA were extracted as described by above, but 65% isopropanol in water (v/v), with BHT (250mg/L) was used as the extraction solvent. From the supernatant, hydrolysis of conjugated IAA and 4-Cl-IAA was carried out using the method described by Symons *et al*. (2002).

For the extraction of Trp and 4-Cl-Trp, tissue was weighed, homogenised and extracted as described above, but distilled water with BHT (250mg/L) was the usual extraction solvent.

3.3 Results and discussion

Determination of the precise evolutionary origin of 4-Cl-IAA in previous study was based on the inability to detect this hormone in a single cultivated species from the *Cicerae* tribe (*Cicer arietinum*), which was taken as a sole representative of the clade immediately basal to the common ancestor of the *Fabeae* and *Trifolieae* (ex. *Parochetus*) tribes (Lam *et al.*, 2015). However, there are a number of genera that are closely related to the *Fabeae* and *Trifolieae*, according to modern phylogenic relationships in the Fabaceae (Wojciechowski *et al.*, 2004; Lavin *et al.*, 2005). The closest of these is the genus *Galega* which is either sister to *Cicer* or immediately basal to *Cicer* (Fig. 4.1). Basal to this clade is the monospecific genus *Parochetus*, paraphyletically long assigned to the *Trifolieae*, followed by the Astragalean clade (for which *Clianthus puniceus* was included as a representative in previous study) (Fig. 3.1).

In order to definitively demonstrate a single evolutionary origin in the common ancestor of the *Fabeae* and *Trifolieae*, (ex. *Parochetus*) tribes dry seed extracts from the perennial species *C. echinospermum* (a wild relative of *C. arietinum*), *Galega officinalis, Parochetus communis* and two additional species within the Astragalean clade (*Astragalus sinicus* and *A. propinquus*) were analysed using base hydrolysis and UPLC-MS/MS as described previously (Symons *et al.*, 2002).

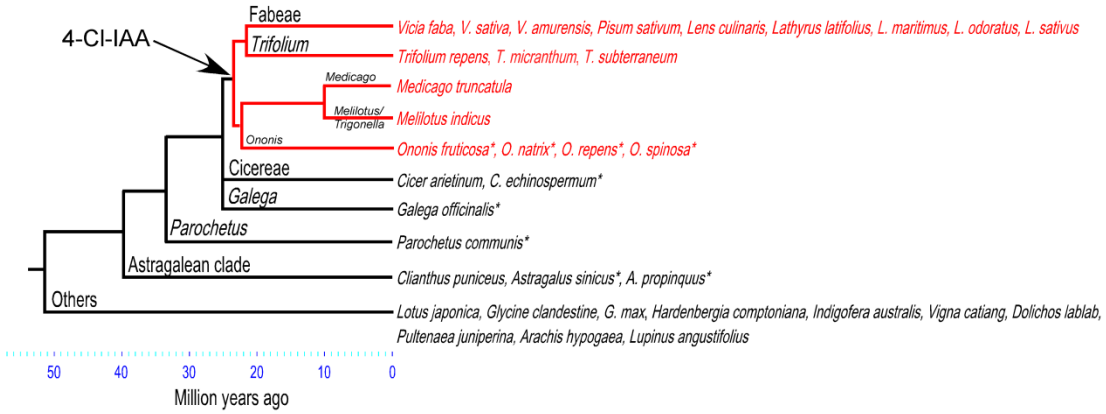


Figure 3.1 Phylogeny of the major lineages of the inverted-repeat-loss clade (IRLC) of the family Fabaceae with a particular focus on the closest lineages to the tribe *Fabeae*. The ability to produce 4-Cl-IAA appears to have evolved after the divergence of the genera *Cicer* and *Galega* approximately 25 million years ago (indicated by arrow and red branches). Species with detectable 4-Cl-IAA are shown in red; species that do not have detectable 4-Cl-IAA are shown in black. Species investigated in this study are indicated by an asterisk. Phylogenetic relationships and divergence dates are taken from Choi *et al.* (2004), Lavin *et al.* (2005), Schaefer *et al.* (2012) and Wojciechowski *et al.* (2004).

UPLC-MS analysis of dry seed extract of *C. echinospermum* revealed that IAA and Trp are present as endogenous compounds (Fig. 3.2). The concentration of IAA was calculated to be 7139 ng.g(FW)$^{-1}$ (Table 3.1). Endogenous 4-Cl-Trp and 4-Cl-IAA were below the limit of detection (Fig. 3.2). Together with *C. arietinum*, this result may indicate that chickpea does not lose chlorinating capacity due to domestication but such ability might not even exist at the very beginning within the genus.

Endogenous IAA and Trp were also detected in *G. officinalis* (Fig. 3.3) as well as *P. communis*, *A. propinquus* and *A. sinicus*. UPLC-MS did not detect the presence of 4-Cl-IAA and 4-Cl-Trp in these species. The total level (free acids and conjugates) of IAA was quantified as 428 ng.g(FW)$^{-1}$, 39235 ng.g(FW)$^{-1}$, 5425 ng.g(FW)$^{-1}$ and 461 ng.g(FW)$^{-1}$ for the species mentioned respectively (Table 3.1).

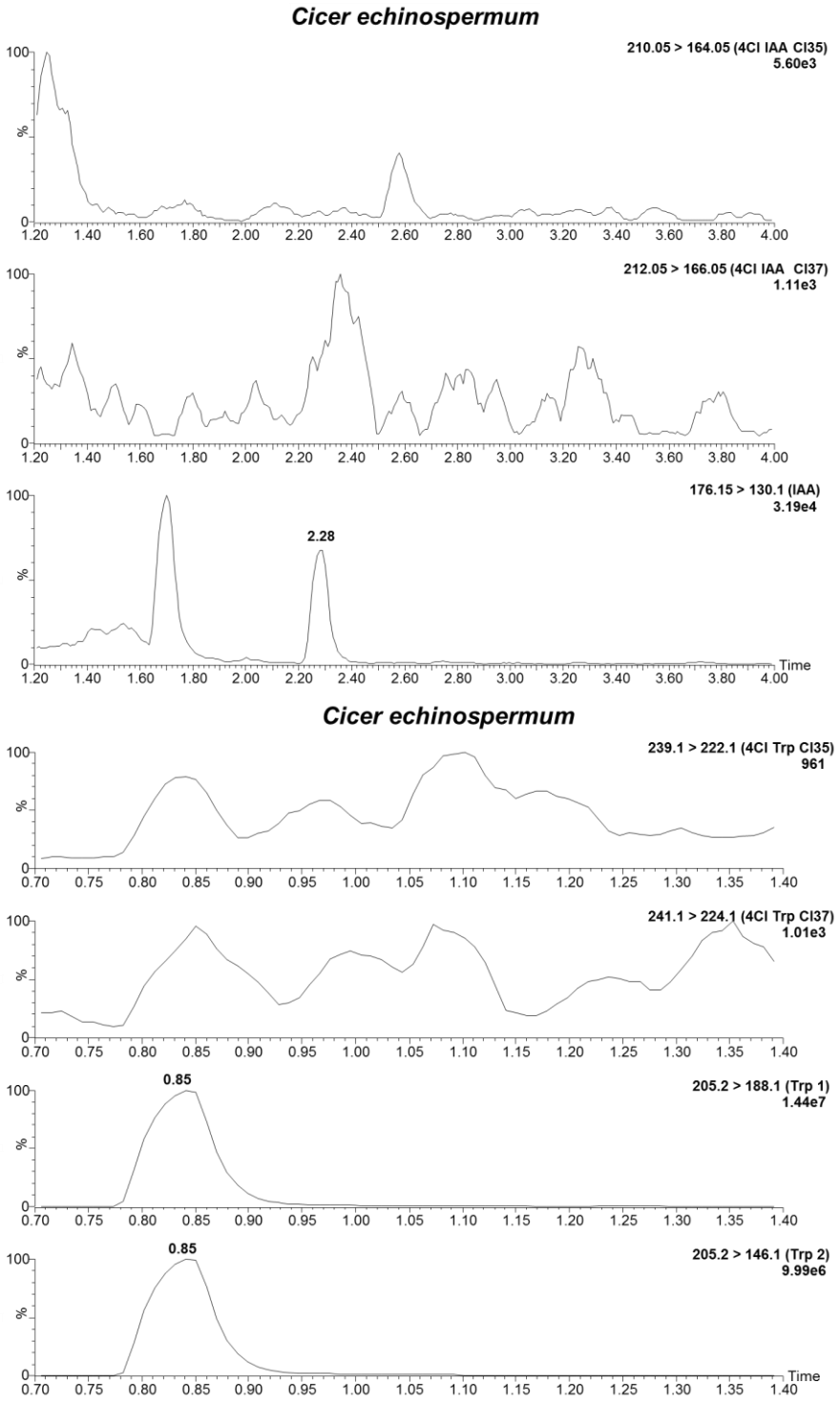


Figure 3.2 UPLC-MS chromatograms (MRM mode) of *Cicer echinospermum* (extract of dry seeds) showing the presence of endogenous IAA (2.28 min) and Trp (0.85 min). 4-Cl-Trp and 4-Cl-IAA were below the limit of detection. Compounds were identified by comparing RTs and MRM transitions with those of relevant standards (data not shown).

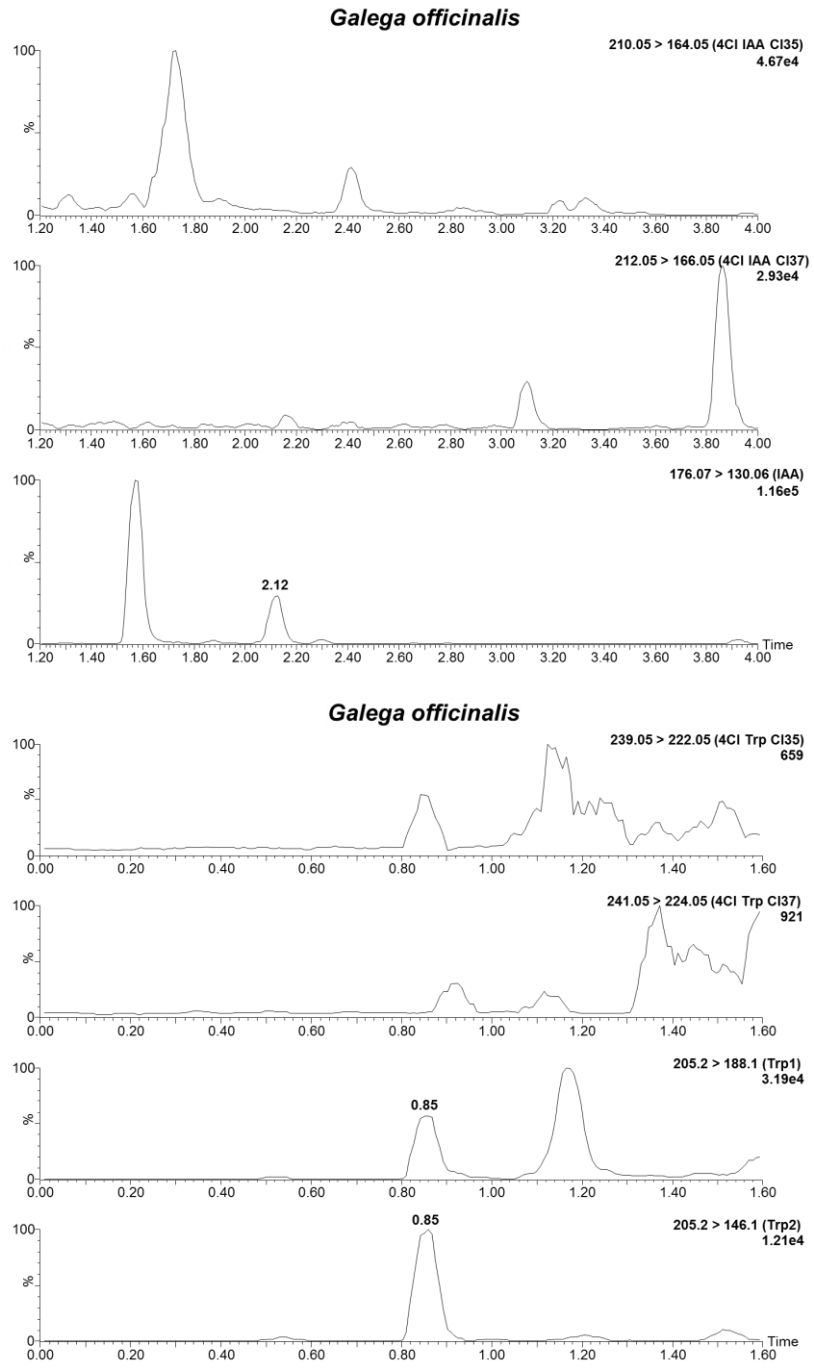


Figure 3.3 UPLC-MS chromatograms (MRM mode) of *Galega officinalis* (extract of dry seeds) showing the presence of endogenous IAA (2.12 min) and Trp (0.85 min). 4-Cl-Trp and 4-Cl-IAA were below the limit of detection. Compounds were identified by comparing RTs and MRM transitions with those of relevant standards (data not shown).

Table 3.1 Total levels (free acids and conjugates) of endogenous 4-Cl-IAA and IAA in 9 representative species from the family Fabaceae. (n.d., not detected)

Species	Phylogenetic clade	Tissue type	IAA ($ng.g^{-1}$)	4-Cl-IAA ($ng.g^{-1}$)
Ononis fruticosa	*Trifolieae*	Dry seeds	3059	1381
Ononis natrix	*Trifolieae*	Dry seeds	3741	3600
Ononis repens	*Trifolieae*	Dry seeds	3555	6261
Ononis spinosa	*Trifolieae*	Dry seeds	2016	4144
Cicer echinospermum	*Cicereae*	Dry seeds	7139	n.d.
Galega officinalis	*Trifolieae*	Dry seeds	428	n.d.
Parochetus communis	*Trifolieae*	Dry seeds	39235	n.d.
Astragalus propinquus	*Astragalae*	Dry seeds	5425	n.d.
Astragalus sinicus	*Astragalae*	Dry seeds	461	n.d.

To confirm the single evolutionary origin of 4-Cl-IAA within the *Fabeae* and *Trifolieae* (ex. *Parochetus*) tribes, four species were sourced from *Ononis*, a genus that may be either basal to the genera *Medicago, Melilotus* and *Trigonella* (Lavin *et al*., 2005; Wojciechowski *et al.*, 2004) or basal to the entire *Fabeae* and *Trifolieae* (ex. Parochetus) clades (Schaefer *et al.*, 2012), and which, as a genus, has not yet been tested for auxins in seeds (Fig. 3.1). Out of sixty-eight described species of *Ononis* (Kloda *et al*., 2008*)*, the species tested here included *O. fruticosa*, *O. natrix*, *O. repens* and *O. spinosa*. Endogenous IAA, Trp, 4-Cl-IAA and 4-Cl-Trp were all identified in the dry seed extracts of all four tested species of *Ononis* in this study (Figs. 3.4 – 3.7). The highest concentration of 4-Cl-IAA, 6261 ng.g(FW)$^{-1}$ was detected in *O. repens* followed by *O. spinosa* (4144 ng.g(FW)$^{-1}$), *O. natrix* (3600 ng.g(FW)$^{-1}$) and *O. fruticosa* (1381ng.g(FW)$^{-1}$) (Table 3.1). As for IAA, *O. natrix* contains the highest level of 3741 ng.g(FW)$^{-1}$ (Table 3.1). *O. repens*, *O. fruticosa* and *O. spinosa* contain 3555 ng.g(FW)$^{-1}$, 3059 ng.g(FW)$^{-1}$ and 2016 ng.g(FW)$^{-1}$ of IAA respectively (Table 3.1). These results show good congruence with previous study and provide stronger evidence for the observation that the chlorinating ability arose along with the divergence of the common ancestor of the *Fabeae* and *Trifolieae* (ex. *Parochetus*) tribes (Fig. 3.1).

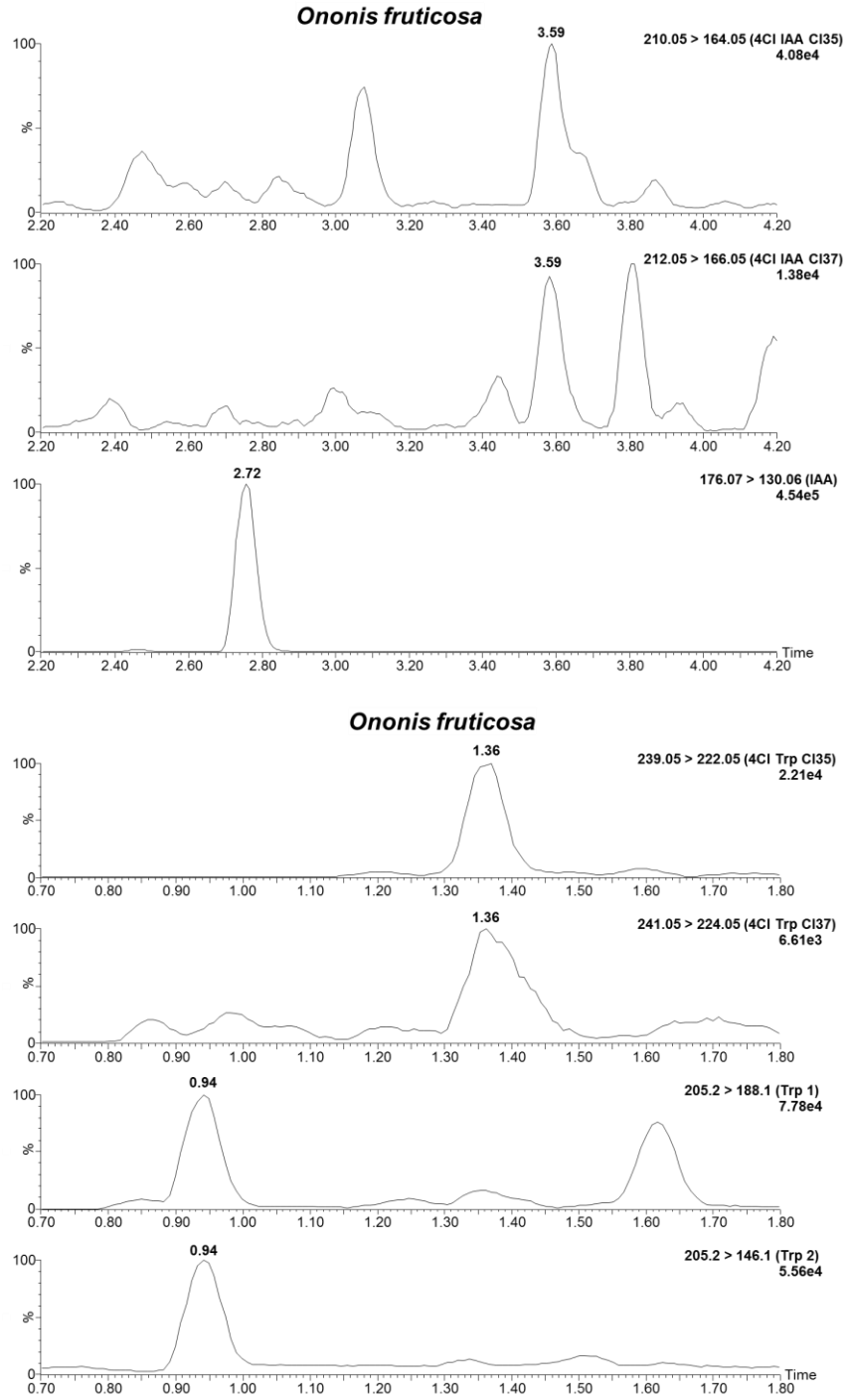


Figure 3.4 UPLC-MS chromatograms (MRM mode) of *Ononis fructicosa* (extract of dry seeds) showing the presence of endogenous 4-Cl-IAA (3.59 min), IAA (2.72 min), 4-Cl-Trp (1.36 min) and Trp (0.85 min). Compounds were identified by comparing RTs and MRM transitions with those of relevant standards (data not shown).

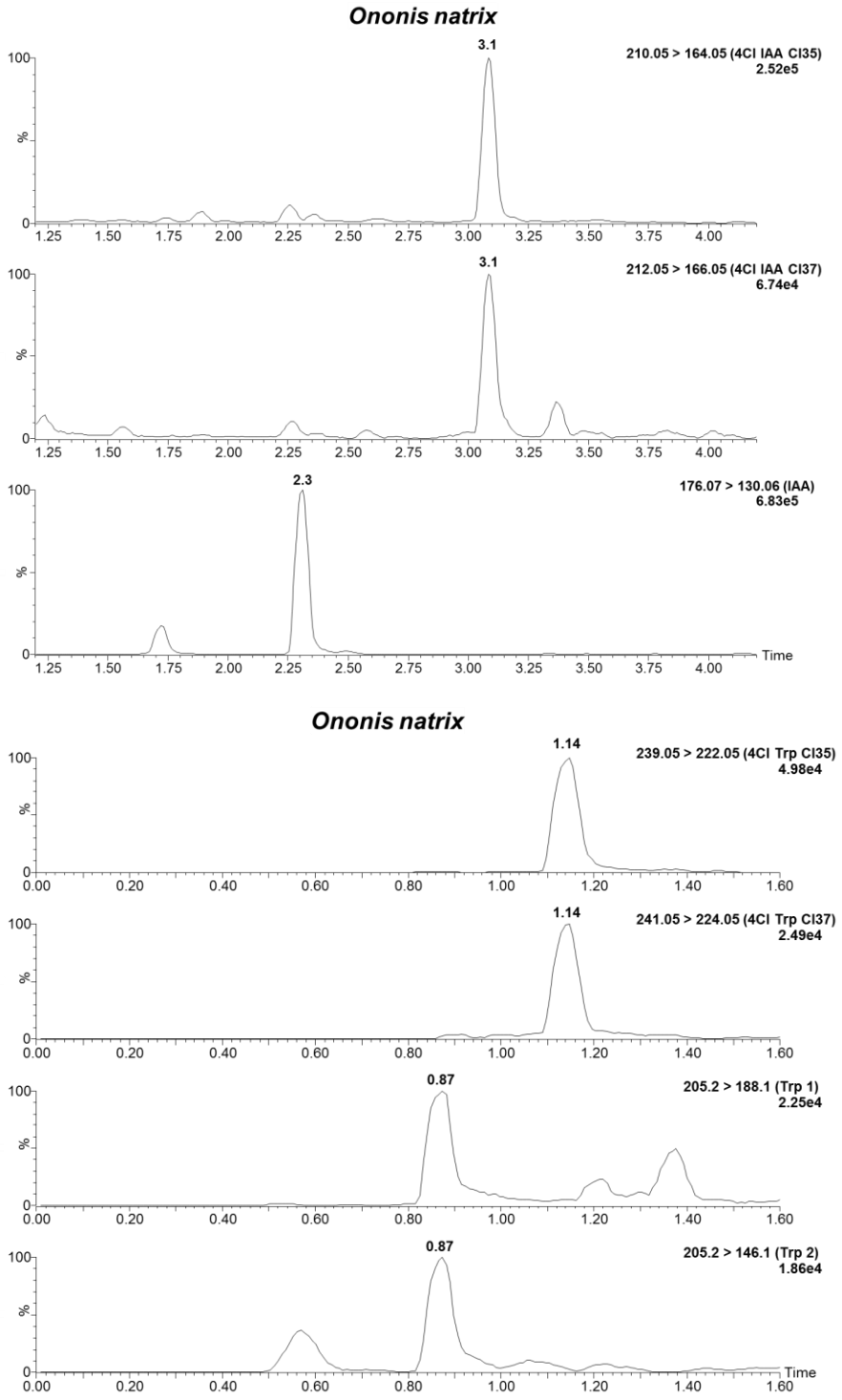


Figure 3.5 UPLC-MS chromatograms (MRM mode) of *Ononis natrix* (extract of dry seeds) showing the presence of endogenous 4-Cl-IAA (3.10 min), IAA (2.30 min), 4-Cl-Trp (1.14 min) and Trp (0.87 min). Compounds were identified by comparing RTs and MRM transitions with those of relevant standards (data not shown).

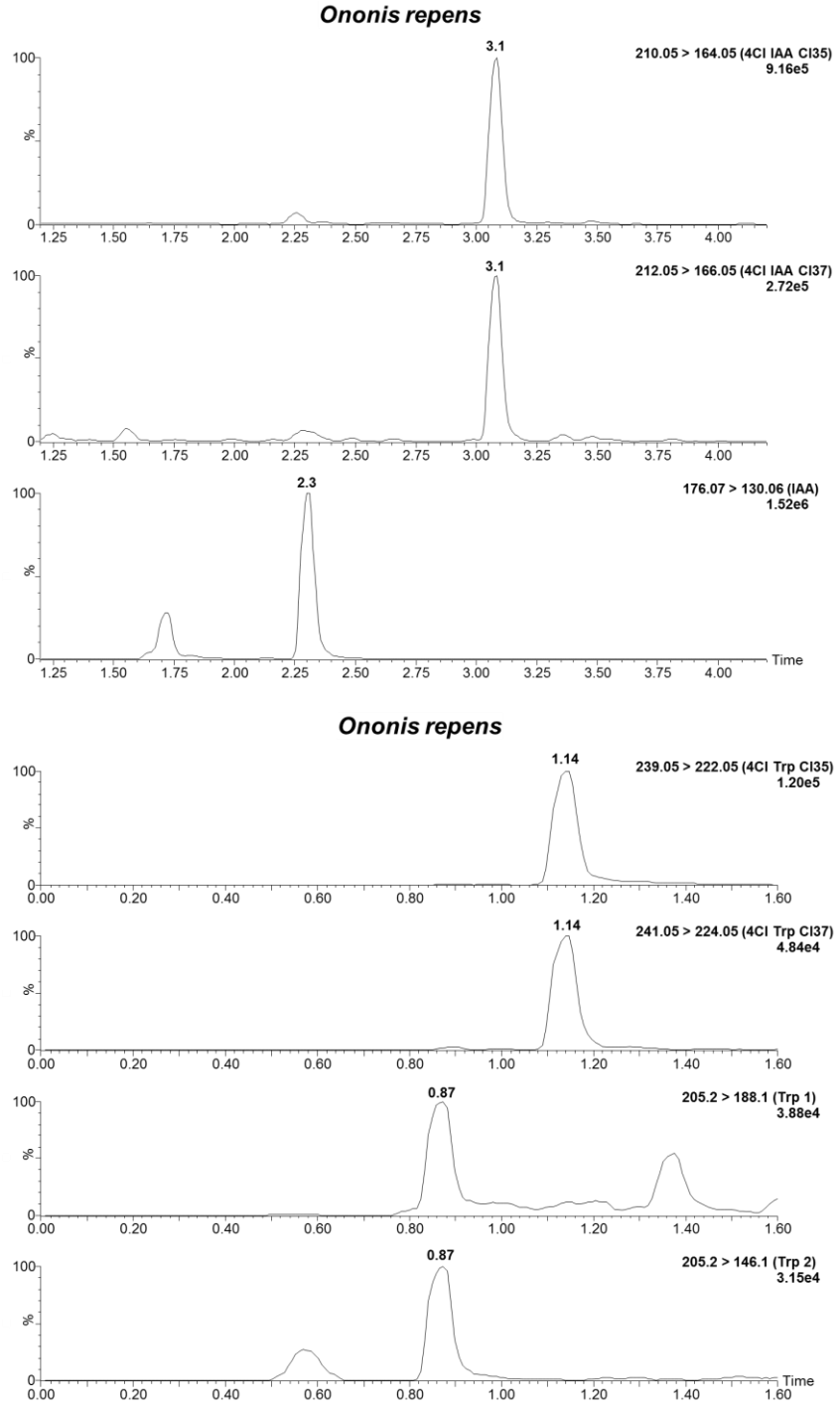


Figure 3.6 UPLC-MS chromatograms (MRM mode) of *Ononis repens* (extract of dry seeds) showing the presence of endogenous 4-Cl-IAA (3.10 min), IAA (2.30 min), 4-Cl-Trp (1.14 min) and Trp (0.87 min). Compounds were identified by comparing RTs and MRM transitions with those of relevant standards (data not shown).

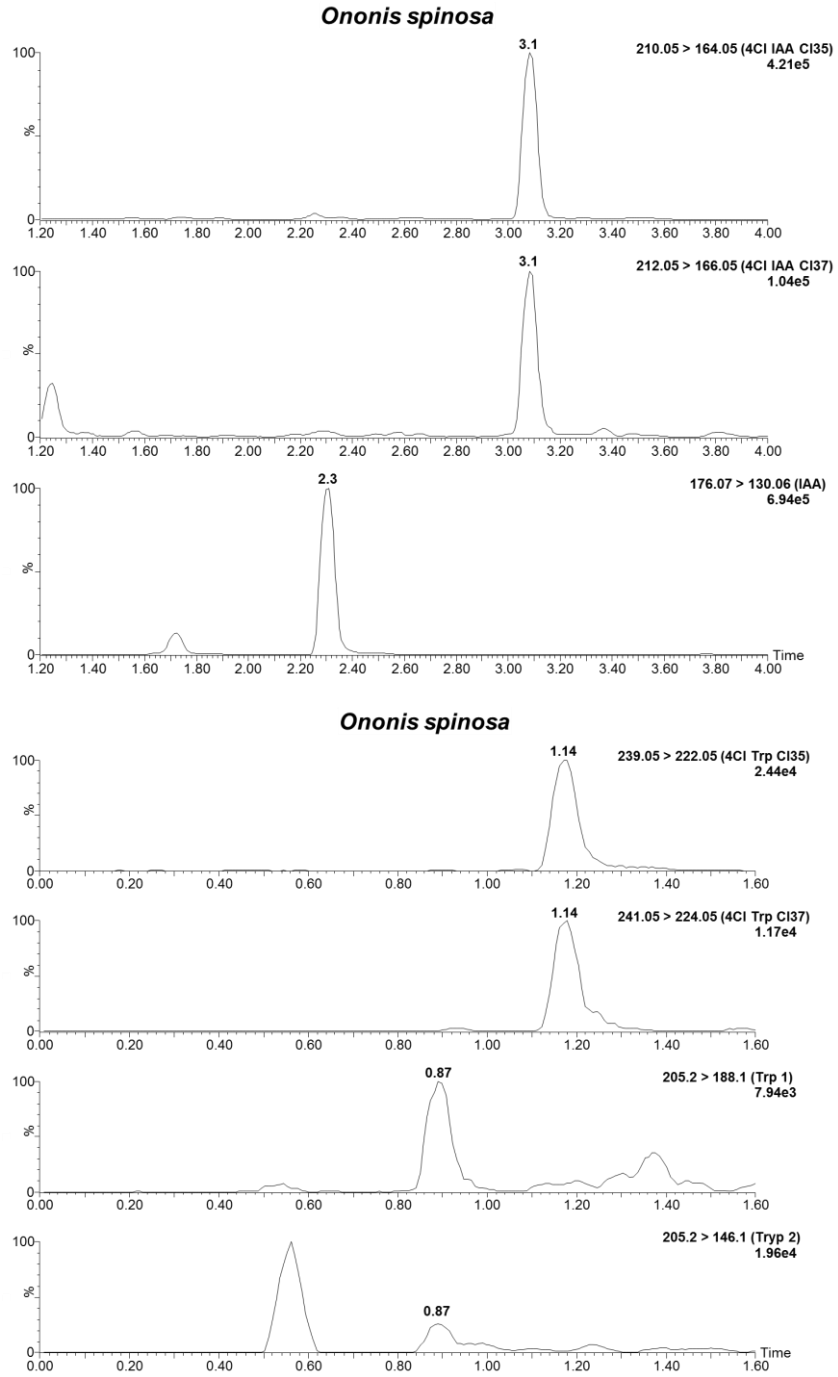


Figure 3.7 UPLC-MS chromatograms (MRM mode) of *Ononis spinosa* (extract of dry seeds) showing the presence of endogenous 4-Cl-IAA (3.10 min), IAA (2.30 min), 4-Cl-Trp (1.14 min) and Trp (0.87 min). Compounds were identified by comparing RTs and MRM transitions with those of relevant standards (data not shown).

A complete survey of Fabaceae (or indeed angiosperm) species containing 4-Cl-IAA would be intriguing, but also highly ambitious as approximately 800 species (from the genera *Pisum*, *Lathyrus*, *Lens*, *Medicago*, *Trifolium*, *Melilotus* and *Vicia*) within this chlorinating clade remain unexamined and species collection, sourcing and routine screening can be difficult. The ability to chlorinate has not yet been examined in species from two remaining genera within the *Fabeae* and *Trifolieae* (ex. *Parochetus*) tribes, namely *Trigonella* and *Vavilovia*. The genus *Trigonella* is sister to the genus *Melilotus* (which has 4-Cl-IAA) and has between 70 and 128 species (Petropoulos, 2002) including fenugreek (*Trigonella foenum-graecum*), an agriculturally important crop and popular spice. In contrast, only a single species comprises the genus *Vavilovia* of the tribe *Fabeae*, *V. formosa*, which is hailed as a relict, endangered species and is basal to the genera *Pisum* and *Lathyrus* (Oskoueiyan *et al.*, 2010; Mikić *et al.*, 2013). Observations of species spanning all of the genera, clades and subgenera within the *Fabeae* and *Trifolieae* (ex. *Parochetus*) tribes will help us to reveal how stable or labile the evolution of chlorinating ability has been in this clade.

The number of leguminous species with known chlorinating capability now stands at 22 and encompasses species from both the *Fabeae* and *Trifolieae* tribes. The data in this study suggest that the ability to chlorinate auxin can be used as a phylogenetically informative trait within the Fabaceae. A pertinent example emerging from this study is the lack of 4-Cl-IAA in the seeds of *Parochetus communis*, a species traditionally assigned to the tribe *Trifolieae*. This study supports the conclusion reached by molecular studies (Lavin *et al.*, 2005; Wojciechowski *et al.*, 2004), that *Parochetus* is not a member of the *Trifolieae* tribe and diverged before the evolution of this clade, as all other *Trifolieae* species so far examined contain 4-Cl-IAA. In conclusion, with the results accrued, we can be increasingly confident that the capacity to produce chlorinated auxin has a single evolutionary origin, most likely in the common ancestor of the *Fabeae* and *Trifolieae* (ex. *Parochetus*) tribes that evolved after the divergence of *Cicer* and *Galega*, 25 million years ago.

Chapter 4 Phylogenetic Relationships in the Fabaceae: Insights from the Possible Loss of Chlorinating Ability in Some Members of the Genus *Trigonella*

4.1 Introduction

The elucidation of a Fabaceae phylogeny has always been challenging even with a number of morphological synapomorphies and burgeoning molecular data made available. In general, the monophyly of subfamily Papilionoideae and at least seven major subclades is well-supported with regard to phylogenetic analysis of plastid *matK rbcL* and *trnL* gene sequence data (Wojciechowski *et al.*, 2004). These informally recognised subclasses include the Cladrastis clade, Genistoid *sensu lato*, Delbergioid *sensu lato*, Mirbelioid, Millettioid, and Robinioid clades, and the inverted-repeat-lacking clade (IRLC) (Wojciechowski *et al.*, 2000; Wojciechowski *et al.*, 2004). The loss of one copy of the inverted repeat in the plastid genome is the defining basis for the IRLC, lacking any strikingly obvious morphological synamorphies (Group, 2013). This most species-rich IRLC subclade is sister to the Robinioids (which include the sequenced species *Lotus japonicus*) (Fig. 4.1) (Wojciechowski *et al.*, 2004). Also, all members of tribes *Carmchaelieae*, *Cicereae*, *Hedysareae*, *Trifolieae*, *Fabeae* (formerly *Viceae*) and *Galegeae* belong to the IRLC (Steele and Wojciechowski, 2003; Wojciechowski *et al.*, 2004). Loss of chloroplast-DNA inverted repeat seems to be non-deleterious but evolutionarily significant because it may have a destabilizing effect on the chloroplast chromosome that in turn gives rise to the comparatively more variable chloroplast DNA of the Leguminosae than in most land plants (Lavin *et al.*, 1990).

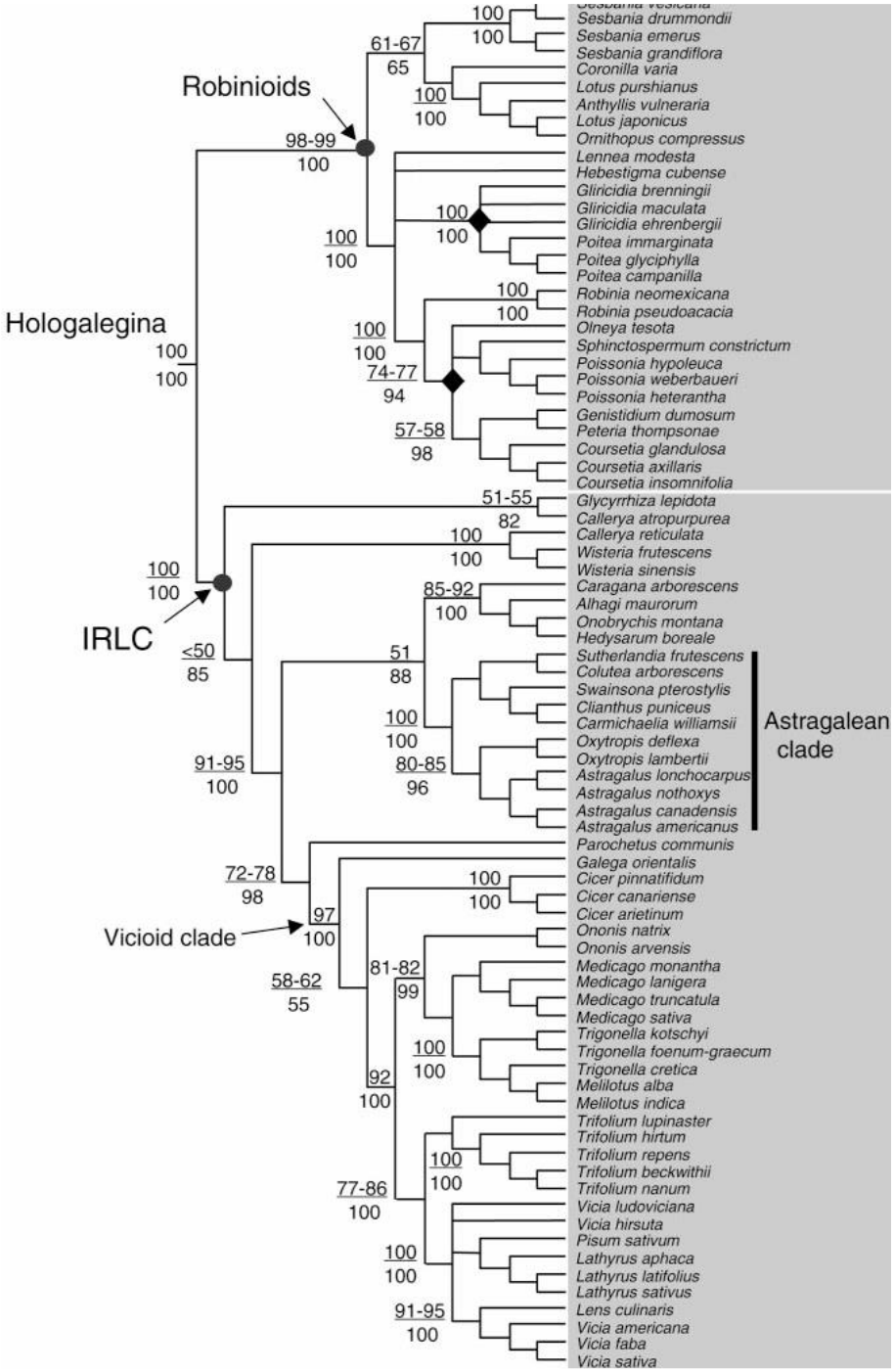


Figure 4.1 Phylogeny of the Robinioid clade and the IRLC. From Wojciechowski *et al.* (2004).

4.1.1 The Vicioids

Within the IRLC, an informally recognised subclade, the Vicioids, is morphologically the most distinctive, comprising many agriculturally important genera including Cicer and the weed genus *Ononis*, the chlorinating genera in tribe *Fabeae* (e.g. *Pisum*, *Lens*, *Lathyrus*, *Vicia*), and tribe *Trifolieae* (e.g. *Melilotus*, *Trigonella*, *Medicago*, *Trifolium*). This generally monophyletic subclade is made up of the tribes *Trifolieae* and *Fabeae* along with *Cicereae* and *Galega* (*Galaegeae*) (Steele and Wojciechowski, 2003). Historically, the tribe *Trigonellinae* comprised *Medicago*, *Melilotus* and *Trigonella* (e.g. in 1901), but was later replaced by the tribe *Trifolieae* erected to include the three genera mentioned and *Trifolium* L. (Steele *et al.*, 2010). *Ononis* and *Parochetus* have also been included by some authors (e.g. Heyn 1981). *Parochetus* is a single-species genus that has been attributed to a range of different tribes (e.g. *Phaseoleae*, *Trifolieae*) or subtribes (e.g. *Phaseolinae*, *Trifoliinae*), based on fruit characters, habit, ovule morphology and seed position within the pods. Distinctive characters of the stipules, leaves, bracts and flowers, taken together with spermoderm pattern, distinguishes it from other *Trifolieae*, and a new, separate subtribe, *Parochetinae*, has been erected (Chaudhary and Sanjappa 1998). It has even been suggested that *Ononis* may deserve a revised tribal or subtribal placement as the morphological characters of the stamens, spermoderm pattern and indumentum of *Ononis* differs from other genera of the *Trifolieae* (Chaudhary and Sanjappa, 1998).

However, *Trifolium* L. may not be sister to the remainder of *Trifolieae* but, instead, to *Fabeae* (Fig. 4.2) (Group, 2013). Despite having unclear relationship to *Trifolieae*, monophyly of *Fabeae* remains well-supported (Steele and Wojciechowski, 2003; Wojciechowski *et al.*, 2004) whereas *Melilotus* Mill. is, either sister to, nested within *Trigonella* L. (Steele and Wojciechowski, 2003; Dangi *et al.*, 2016).

Medicago includes 23 species that were previously assigned to *Trigonella* (the so-called "medicagoid *Trigonella*" which currently comprises sections *Bucerates* and *Lunatae*) transferred by Small *et al.* (1987) based on floral features associated with the explosive pollination syndrome. However, it is generally resolved as monophyletic and sister to *Trigonella* (Steele and Wojciechowski, 2003).

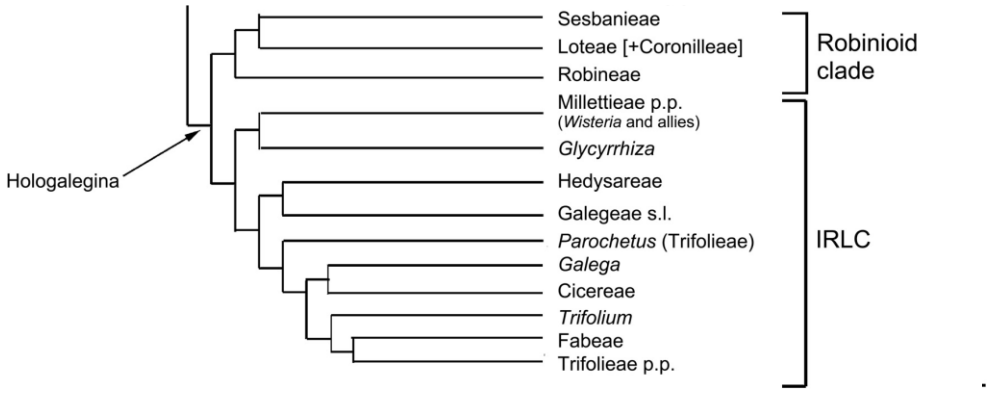


Figure 4.2 Schematic phylogeny of Papilionoideae compiled as a supertree based upon phylogenetic analyses.. From Group (2013).

4.1.2 *Trigonella*

Bayesian inference of combined data and trees resulting from maximum parsimony make *Trigonella* an evidently monophyletic genus, disagreeing with the traditional subgeneric division (Dangi *et al.*, 2016). Morphology is still the major key to sectional delimitation among species in *Trigonella* (Sirjaev, 1928-1934). In particular, the sectional delimitation of sections *Cylindricae*, *Ve´rae*, *Samaroideae*, *Pectinatae*, *Erosae* and *Callicerates* is supported by phylograms and by strong support for combining sections *Erosae* and *Pectinatae* into one; apomorphies in inflorescence and stipule supported three major lineages as indicated by phylogenetic reconstructions (Dangi *et al.*, 2016). In contrast to their current classification, species from sections *Foenum-graecum* and *Falcatulae* clustered in different subclades (Dangi *et al.*, 2016). In fact, within genus *Trigonella* at subgeneric and sectional level, inflorescence type provides taxonomic potential for phenetic implications. Some taxonomically important legume and seed characters are thought to be homeoplastic and have arisen more than once in *Trigonella*.

4.1.3 Resolving taxonomic placement: choice of useful characteristics

In a phylogentic context, the choice of characteristics employed, (morphological, biochemical or molecular) is very important and in many cases dictates how powerful and meaningful the outcome is. For instance, total evidence analyses, optimising morphological characters onto a molecular phylogeny or consideration of morphological data in a new perspective will reveal morphological snynamorphies for clades and hence provide new phylogenetic classifications, if not all, as well as many evolutionary insights (Group, 2013). In other words, the addition of morphological data to a gene sequence dataset can provide greater resolution and stronger evidence for clade support in the resulting phylogeny (Group, 2013).

There are also morphological characters not readily visible to the naked eye; one example stems from a unique class of ATP-independent contractile proteins called forisomes that can plug and reopen sieve tubes, regulating phloem transport (Peters *et al.*, 2010). All studied IRLC species have either tailless forisomes or lack forisomes, other than two exceptions (for more details refer to Peters *et al*. 2010).

Mostly found in the Astragalean clade, the tribe Galegeae has forisome-lacking genera except for *Galega* (part of the Vicioid clade) (Fig. 4.3) (Peters *et al.*, 2010). The species in the Vicioid clade only include species with tailless forisomes except for *Galega* which does not possess forisomes (Peters *et al.*, 2010). Yet there are two types of forisomes: with or without tail-like protrusions, of which the occurrence of the former is less frequent. The forisome type, however, can be rendered as a reliable criterion for classification.

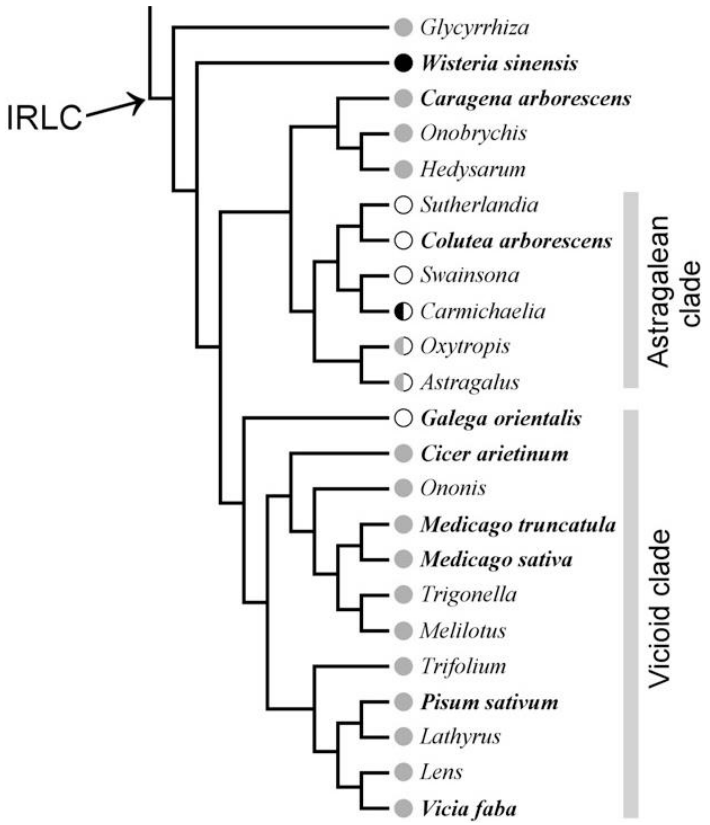


Figure 4.3 Distribution of tailed and tailless forisomes mapped onto the molecular phylogenetic tree of the *Papilionoideae* of Wojciechowski *et al*. (2004), whose nomenclature for taxa above the genus level is followed in this graph. Only genera and species that were represented in that phylogenetic tree as well as in the data set in this study are shown. The tree drawn serves to visualize the topology of the clades; branch lengths do not represent phylogenetic distances. Circles in front of taxon names stand for tailed forisomes (black circles), tailless forisomes (gray circles), or absence of forisomes (open circles). IRLC, inverted repeat-lacking clade. From Peters *et al.* (2010).

In terms of molecular characteristics, delimitation of *Medicago* and *Trigonella* as sister genera has always been strongly supported, based on analyses of both the nrDNA, ITS and the flanking nrDNA ETS, as well as the plastid-encoded matK gene (Steele *et al.*, 2010).

Another example involves using biochemical characteristics to further differentiate *Medicago, Trigonella* and *Melilotus*. For instance, species of *Medicago* accumulate vesitol and sativan (the type of post-fungal infection phytoalexins), and contain haemolytic saponins, while these substances are absent in *Trigonella* or *Melilotus* (Steele *et al.*, 2010).

4-Cl-IAA and its chlorinated precursors are deemed to be hallmarks of specific genera and tribes within the Fabaceae. Here I explore the possibility that the capacity to chlorinate is a useful characteristic that may aid unravelling the taxonomic relationships in legumes. This chapter will report the intriguing observation of non-chlorinating *Trigonella* species. This observation raises the important question of how universal is possible chlorinating ability within the *Fabeae-Trifolie*ae, this will be answered by further screening of the other previously known chlorinating genera.

4.2 Materials and methods

4.2.1 Plant materials

Dry seeds of 28 *Trigonella* species were sourced from Australian Pasture Genebank (APG). In particular, *T. macrorrhyncha*, *T. glabra*, *T. schlumbergeri*, *T. cretica* and *T. spicata* were grown in glasshouse conditions as described previously (Jager *et al.*, 2007) and their young seeds or pods were harvested for hormonal analysis. In addition, 15 *Melilotus* species (comprising the entire genus), 10 *Medicago* species (spanning all major tribes), 4 *Ononis* species and 3 *Trifolium* species surveyed in this study were obtained from APG.

4.2.2 Extract preparation for the detection and quantification of compounds

For the extraction and quantification of IAA and 4-Cl-IAA from young, fresh tissues, 0.3-2.5g of tissue was weighed (±0.0001 g FW), placed into a falcon tube with four volumes of cold (-20°C) extraction solvent (80% methanol in water (v/v), with BHT (250 mg/L). The tissue was then homogenised and held at 4°C overnight to extract. For each species, the supernatant was then divided in half, in order to conduct two separate analyses: one for detection of the compounds of interest, and one for quantification of these compounds. For the former, no labelled internal standards were added; and for the latter, [$^{13}C_6$] IAA (Cambridge Isotope Laboratories) and [D_2] 4-Cl-IAA (supplied by Prof. Jerry Cohen, Department of Horticultural Science, University of Minnesota) were added as internal standards. The samples were reduced under vacuum at 35°C, taken up in 2% acetic acid in water (v/v), and partitioned twice against diethyl ether. After drying the ether, samples were taken up in 1% acetic acid in water (v/v) and centrifuged for 5 min at 13 000 g. Aliquots were then taken for analysis by UPLC-MS/MS as described below previously (Tivendale *et al.*, 2012).

In viable, dry seeds, the total levels of each of IAA and 4-Cl-IAA were monitored, including both free acids and conjugated forms. IAA and 4-Cl-IAA were extracted as described by above, but 65% isopropanol in water (v/v), with BHT (250mg/L) was used as the extraction solvent. From the supernatant, hydrolysis of conjugated IAA and 4-Cl-IAA was carried out using the method described by Symons *et al.* (2002).

For the extraction of Trp and 4-Cl-Trp, tissue was weighed, homogenised and extracted as described above, but distilled water with BHT (250mg/L) was the usual extraction solvent.

4.3 Results and discussion

In this study, 28 species representing all 13 sections recognised in *Trigonella* (Sirjaev, 1928-1934; Ranjbar and Hajmoradi, 2015) were screened for 4-Cl-IAA and 4-Cl-Trp (Table 4.1). Most of the *Trigonella* species tested were samples extracted from dry seeds using base hydrolysis. This method actually reveals the total 4-Cl-IAA (including 4-Cl-IAA in the form of methyl ester that seems to be abundant in dry mature seeds) produced or stored over the full course of seed development. 17 out of 27 *Trigonella* species showed evidence of detectable 4-Cl-IAA and 4-Cl-Trp. Three *Trigonella* species, including *T. foenum-graecum* (section *Foenum-graecum*) (Fig. 4.4), *T. gladiata* (also section *Foenum-graecum*) (Fig. 4.5) and *T. spicata* (section *Uncinatae*) (Fig. 4.6) did not contain detectable levels of either chlorinated Trp or chlorinated IAA. The chlorinating ability of *T. balansae* and *T. anguina* (section *Falcutulae*) as well as *T. calliceras* (section *Callicerates*), *T. strangulata* (section *Cylindricae*), *T. coelesyriaca* (section *Verae*), *T. spinosa* (section *Spinosae*), *T. caerulea* (section *Capitatae*) and *T. coerulescens* (section *Foenum-graecum*) were also questionable with 4-Cl-Trp in these species detected but not 4-Cl-IAA (Fig. 4.7) (Table 4.1).

Dangi *et al*. (2016) reported a strongly supported clade in section *Foenum-graecum* formed by *T. foenum-graecum* and *T. gladiata* (Fig. 4.8) of which both were found here to be not producing 4-Cl-IAA and 4-Cl-Trp. These two species are sister to a strongly supported subclade including *T. balansae* and *T. anguina* (both traditionally assigned to section *Falcatulae*), both of which were also found to have no detectable 4-Cl-IAA in dry seeds. However, the latter two species did produce chlorinated Trp, suggesting chlorinating ability per se is not lost, just the ability to synthesise 4-Cl-IAA from 4-Cl-Trp. This clade of four species found to lack 4-Cl-IAA are sister to a group that includes a clade of the species *T. spicata* and *T. strangulata*, both species being assigned to section *Uncinatae* and found here to also neither contain 4-Cl-IAA in either fresh pods or dry seeds, respectively (Table 4.1) (Fig. 4.10). While these species are sister to section *Cylindricae* of which all 6 species examined have detectable 4-Cl-IAA and 4-Cl-Trp in seeds, this entire clade is sister to another clade including *T. coerulescens* and *T. caerulea* both of which also do not have detectable 4-Cl-IAA in dry seeds (Fig. 4.10). The absence of 4-Cl-IAA in so many closely

related species suggests a common, single-origin loss (and subsequent reversion) of the ability to synthesise 4-Cl-IAA occurring in the common ancestor of these species (as delineated by the red lines in the ancestors of this clade, Fig. 4.8). In addition to these closely related species lacking 4-Cl-IAA, two other species scattered across the phylogeny of *Trigonella* were found to not contain 4-Cl-IAA, *T. calliceras* and *T. spinosa*.

Based on research in *Arabidopsis thaliana*, *Pisum satium* and *Zea mays*, the two-step linear IAA biosynthetic pathway is comprised of the conversion of Trp to IPya by TAA1/TAR enzymes and of IPyA to IAA by YUCs (Tivendale *et al.* 2014). Parallel IPyA- and 4-Cl-IPyA pathways have been proposed based on several labelled-precursors feeding studies in pea seeds (Tivendale *et al.* 2012). The detection of 4-Cl-Trp in the seeds of those non-4-Cl-IAA producing *Trigonella* species may indicate that the TAR enzymes failed to convert 4-Cl-Trp to 4-Cl-IPyA and/or the YUCs failed to mediate the subsequent conversion of 4-IPyA to 4-Cl-IAA. The unknown Trp-halogenating enzyme is, however, still functioning in these species. The accumulation of IAA in the seeds of these species also indicates that certain TARs and YUCs are functional in the conversion of Trp to IPyA and IPyA to IAA. Isolation and functional assay of *Trigonella* TAA1/TAR enzymes in the future will be valuable for the understanding substrate specificity in these proteins.

In fact, retention time and fragmentation patterns are the two most significant bases for identification using MS, whereas the approximate 3:1 ratio of 35Cl:37Cl is also diagnostic for chlorinated compounds. For 4-Cl-Trp, when the signals are strong (e.g. Fig. 4.9), there is indeed a 3:1 ratio (approximately) but in other cases, where perhaps the signal is not so strong, the ratio was slightly less than 3:1. This reflects the variation that might be expected with weaker samples.

Table 4.1 Endogenous levels of 4-Cl-IAA and IAA in 28 representative species from the genus *Trigonella*. (n.d. not detected).

Subgenus/section	Species	Tissue type	IAA ($ng.g^{-1}$)	4-Cl-IAA ($ng.g^{-1}$)	4Cl-Trp
Subgenus I: *Trigonella*					
Falcatulae	*Trigonella anguina*	Dry seeds	6163	n.d.	Y
	Trigonella balansae	Dry seeds	108	n.d.	Y
	Trigonella maritima	Dry seeds	51	25	Y
	Trigonella stellata	Dry seeds	284	306	Y
	Trigonella suavissima	Dry seeds	282	202	Y
	Trigonella ornithopodioides	Dry seeds	116	810	Y
Callicerates	*Trigonella calliceras*	Dry seeds	29	n.d.	Y
Uncinatae	*Trigonella spicata*	Fresh pods	48	n.d.	N
Cylindricae	*Trigonella cylindracea*	Dry seeds	79	73	Y
	Trigonella filipes	Dry seeds	110	225	Y
	Trigonella hierosolymitana	Dry seeds	77	59	Y
	Trigonella kotschyi	Dry seeds	548	381	Y
	Trigonella mesopotamica	Dry seeds	1454	623	Y
	Trigonella strangulata	Dry seeds	110	n.d.	Y
	Trigonella	Dry seeds	621	1297	Y

	spruneriana				
Samaroideae	*Trigonella cretica*	Fresh pods	25	19	Y
Pectinatae	*Trigonella arabica*	Dry seeds	219	381	Y
Erosae	*Trigonella schlumbergeri*	Fresh pods	20	54	Y
Verae	*Trigonella coelesyriaca*	Dry seeds	18	n.d.	Y
	Trigonella grandiflora	Dry seeds	35	106	Y
Spinosae	*Trigonella spinosa*	Dry seeds	14	n.d.	Y
Bucerates	*Trigonella geminiflora*	Dry seeds	68	58	Y
	Trigonella glabra	Young seeds	326	267	Y
Subgenus II: *Trifoliastrum*					
Capitatae	*Trigonella caerulea*	Dry seeds	110	n.d.	Y
Subgenus III: *Foenum-graecum*					
Foenum-graecum	*Trigonella foenum-graecum*	Dry seeds	10	n.d.	N
	Trigonella gladiata	Dry seeds	33	n.d.	N
	Trigonella macrorrhyncha	Fresh pods	3	43	Y
	Trigonella coerulescens	Dry seeds	42	n.d.	Y

Subgenus and sections in *Trigonella* were referred to Martin *et al.* (2011), Dangi *et al.* (2016), Ranjbar and Zahra (2016) and Sirjaev (1928-1934)

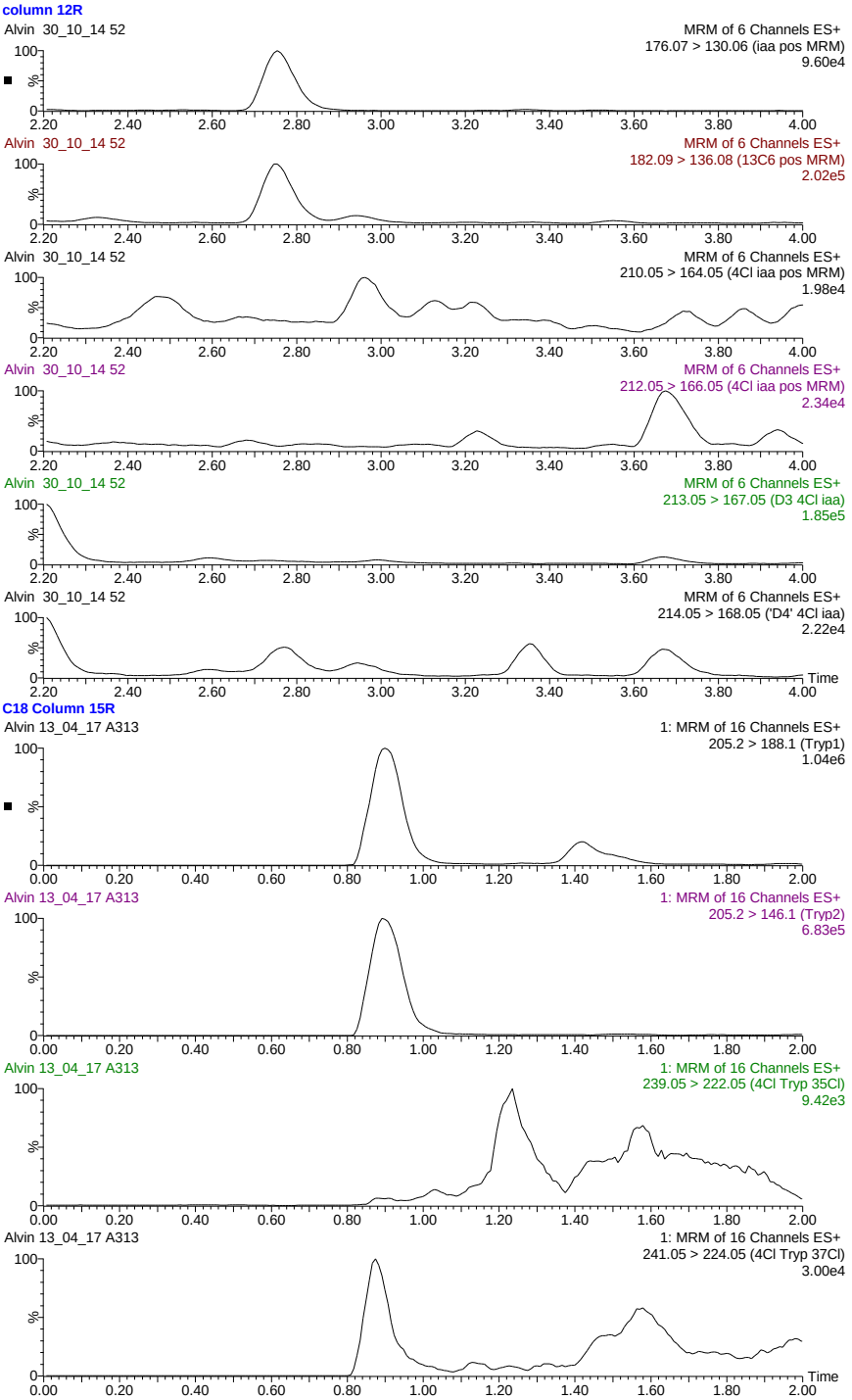


Figure 4.4 UPLC-MS chromatograms (MRM mode) of *Trigonella foenum-graecum* (extract of dry seeds) showing the presence of endogenous IAA (2.76 min) and Trp (0.9 min). 4-Cl-Trp (1.1 min) and 4-Cl-IAA (3.68 min) were below the limit of detection. Peak intensities (normalised) are plotted against retention time; the height of the strongest peak is shown at the right. Transitions: endogenous IAA, 176-130; [$^{13}C_6$]-IAA, 182-136; endogenous 4-Cl-IAA, 210-164; [$^{2}H_2$]4-Cl-IAA, 212-166; [$^{2}H_3$]4-Cl-IAA, 213-167; [$^{2}H_4$]4-Cl-IAA, 214-168.

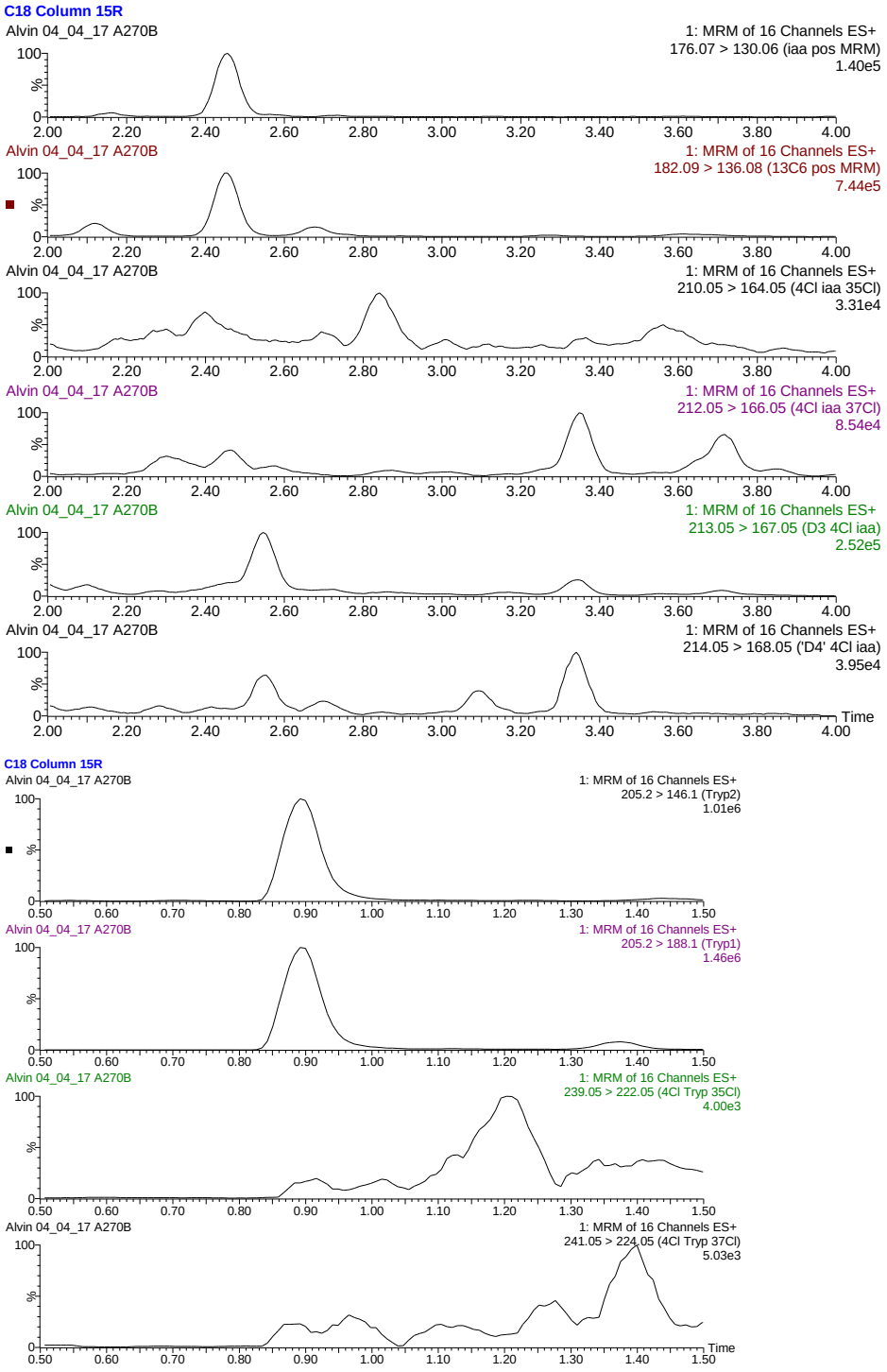


Figure 4.5 UPLC-MS chromatograms (MRM mode) of *Trigonella gladiata* (extract of dry seeds) showing the presence of endogenous IAA (2.46 min) and Trp (0.9 min). 4-Cl-Trp (1.1 min) and 4-Cl-IAA (3.24 min) were below the limit of detection. Peak intensities (normalised) are plotted against retention time; the height of the strongest peak is shown at the right. Transitions: endogenous IAA, 176-130; [$^{13}C_6$]-IAA, 182-136; endogenous 4-Cl-IAA, 210-164; [$^{2}H_2$]4-Cl-IAA, 212-166; [$^{2}H_3$]4-Cl-IAA, 213-167; [$^{2}H_4$]4-Cl-IAA, 214-168.

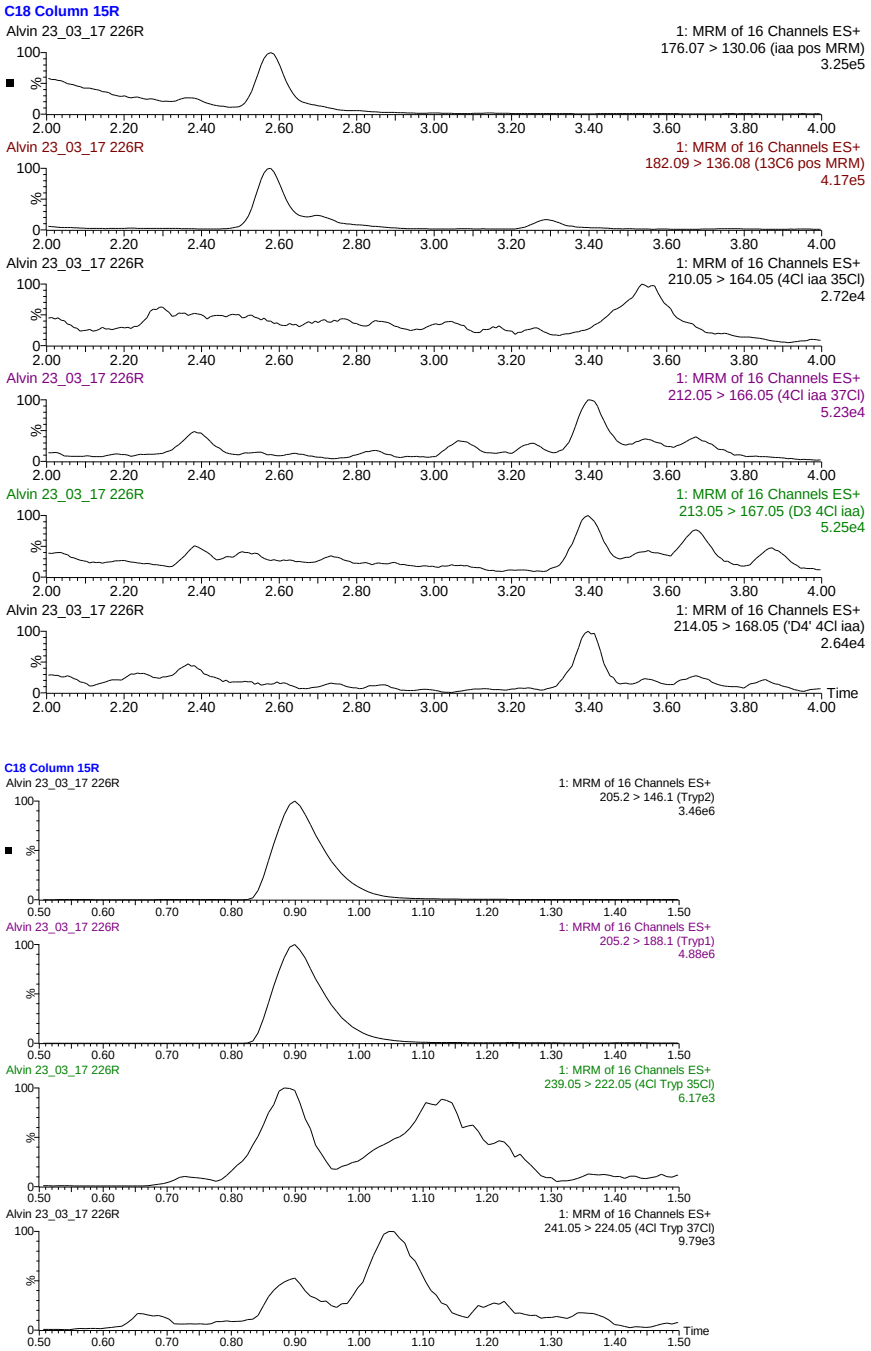


Figure 4.6 UPLC-MS chromatograms (MRM mode) of *Trigonella spicata* (extract of dry seeds) showing the presence of endogenous IAA (2.58 min) and Trp (0.9 min). 4-Cl-Trp (1.1 min) and 4-Cl-IAA (3.40 min) were below the limit of detection. Peak intensities (normalised) are plotted against retention time; the height of the strongest peak is shown at the right. Transitions: endogenous IAA, 176-130; [$^{13}C_6$]-IAA, 182-136; endogenous 4-Cl-IAA, 210-164; [$^{2}H_2$]4-Cl-IAA, 212-166; [$^{2}H_3$]4-Cl-IAA, 213-167; [$^{2}H_4$]4-Cl-IAA, 214-168.

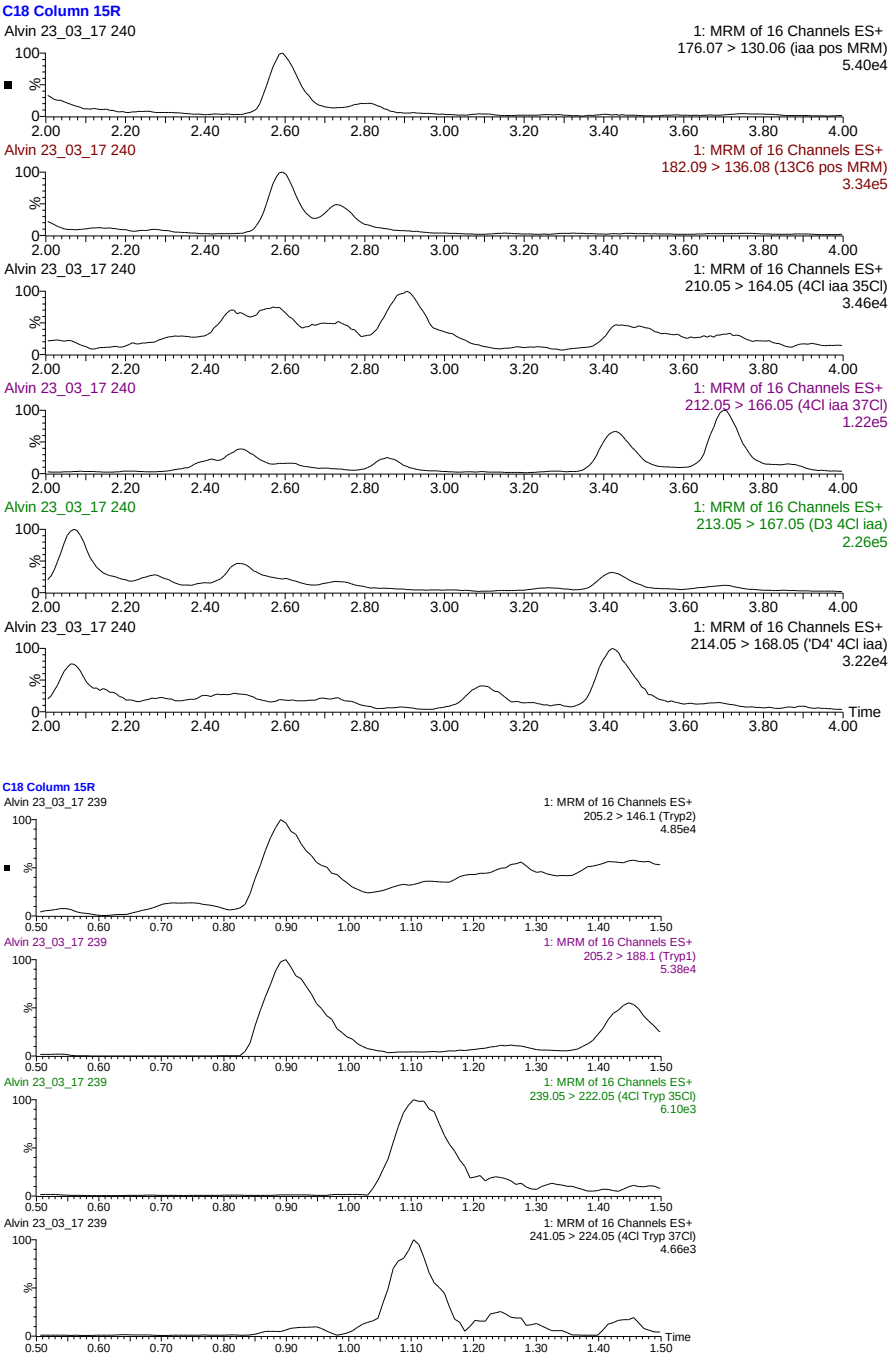


Figure 4.7 UPLC-MS chromatograms (MRM mode) of *Trigonella balansae* (extract of dry seeds) showing the presence of endogenous IAA (2.60 min) and Trp (0.9 min). Limited evidence of 4-Cl-IAA (3.40 min) whereas signals of 4-Cl-Trp (1.1 min) do not conform with 3:1 ratio of ^{35}Cl 4-Cl-Trp to ^{37}Cl 4-Cl-Trp. Peak intensities (normalised) are plotted against retention time; the height of the strongest peak is shown at the right. Transitions: endogenous IAA, 176-130; [$^{13}C_6$]-IAA, 182-136; endogenous 4-Cl-IAA, 210-164; [$^{2}H_2$]4-Cl-IAA, 212-166; [$^{2}H_3$]4-Cl-IAA, 213-167; [$^{2}H_4$]4-Cl-IAA, 214-168.

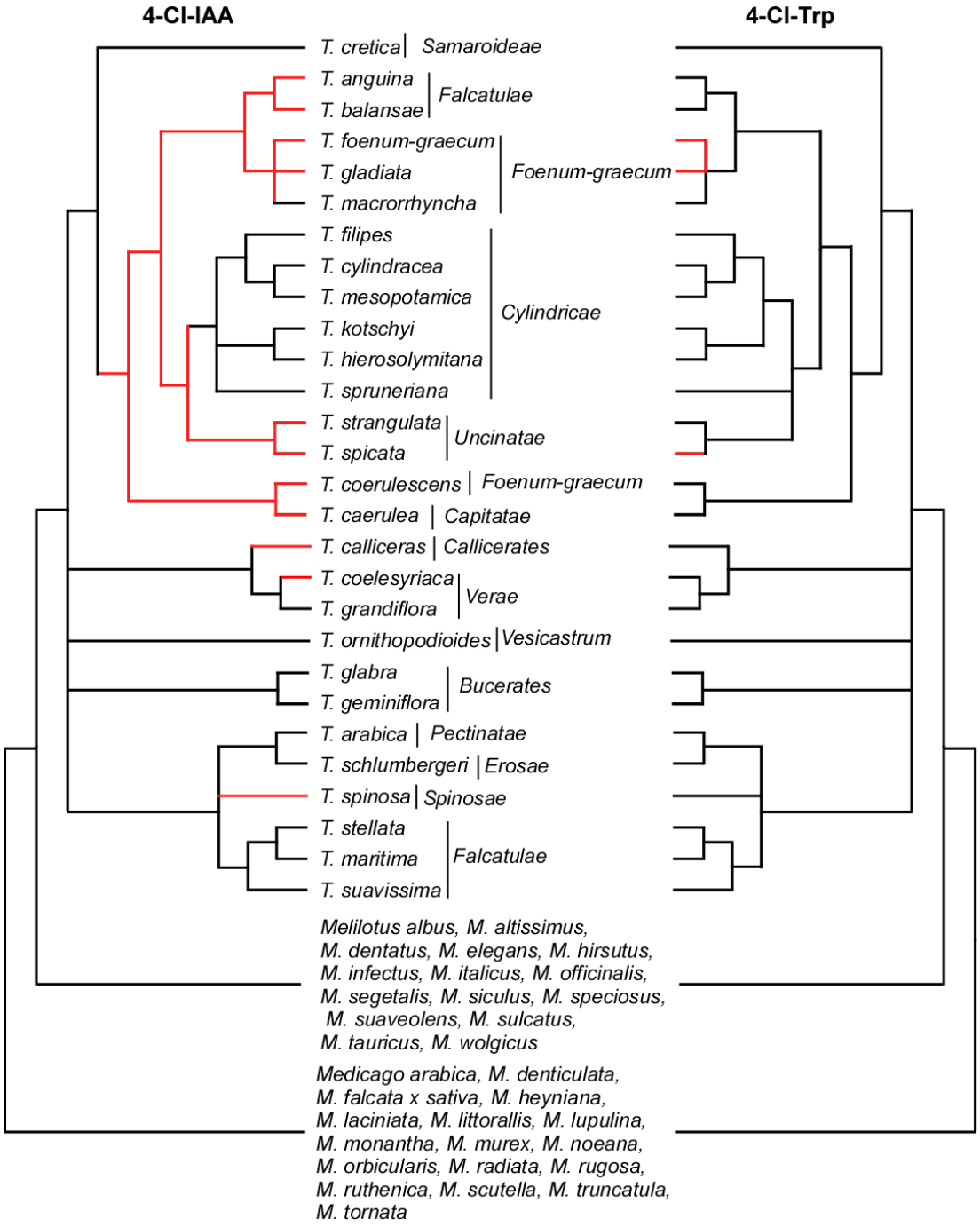


Figure 4.8 Phylogenetic tree of *Trigonella*, *Melilotus* and *Medicago* based on maximum parsimony analysis adapted from Dangi *et al*. (2016) in relation to the data in this study documenting the presence or absence of 4-Cl-IAA and 4-Cl-Trp in the dry seeds of the species. Red branches indicate loss/absence of either 4-Cl-IAA or 4-Cl-Trp while black branches indicate the presence of the compound of interest.

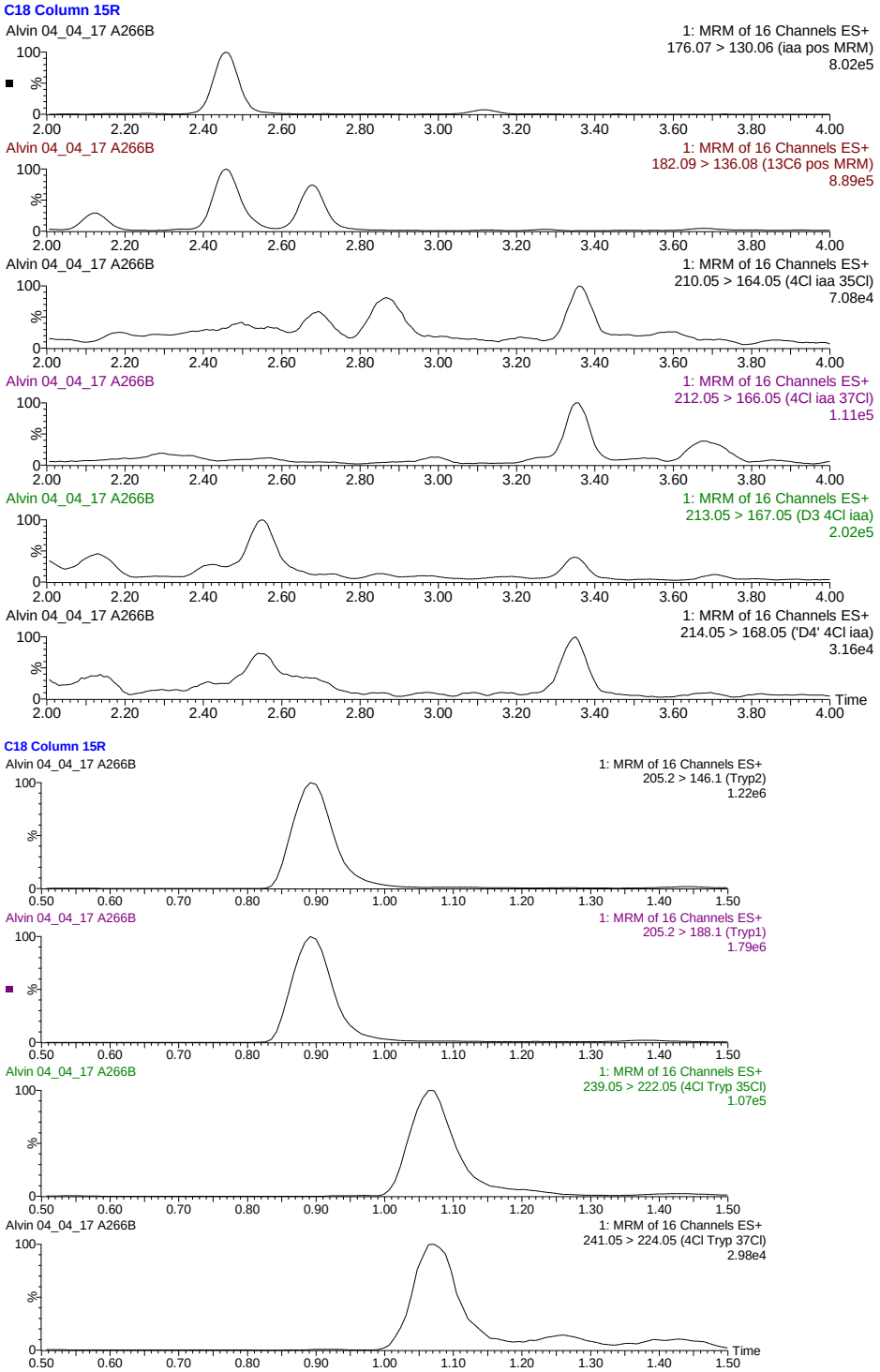


Figure 4.9 UPLC-MS chromatograms (MRM mode) of *Trigonella cylindracea* (extract of dry seeds) showing the presence of endogenous IAA (2.46 min), 4-Cl-IAA (3.36 min), Trp (0.9 min) and 4-Cl-Trp (1.08 min). Peak intensities (normalised) are plotted against retention time; the height of the strongest peak is shown at the right. Transitions: endogenous IAA, 176-130; [$^{13}C_6$]-IAA, 182-136; endogenous 4-Cl-IAA, 210-164; [$^{2}H_2$]4-Cl-IAA, 212-166; [$^{2}H_3$]4-Cl-IAA, 213-167; [$^{2}H_4$]4-Cl-IAA, 214-168.

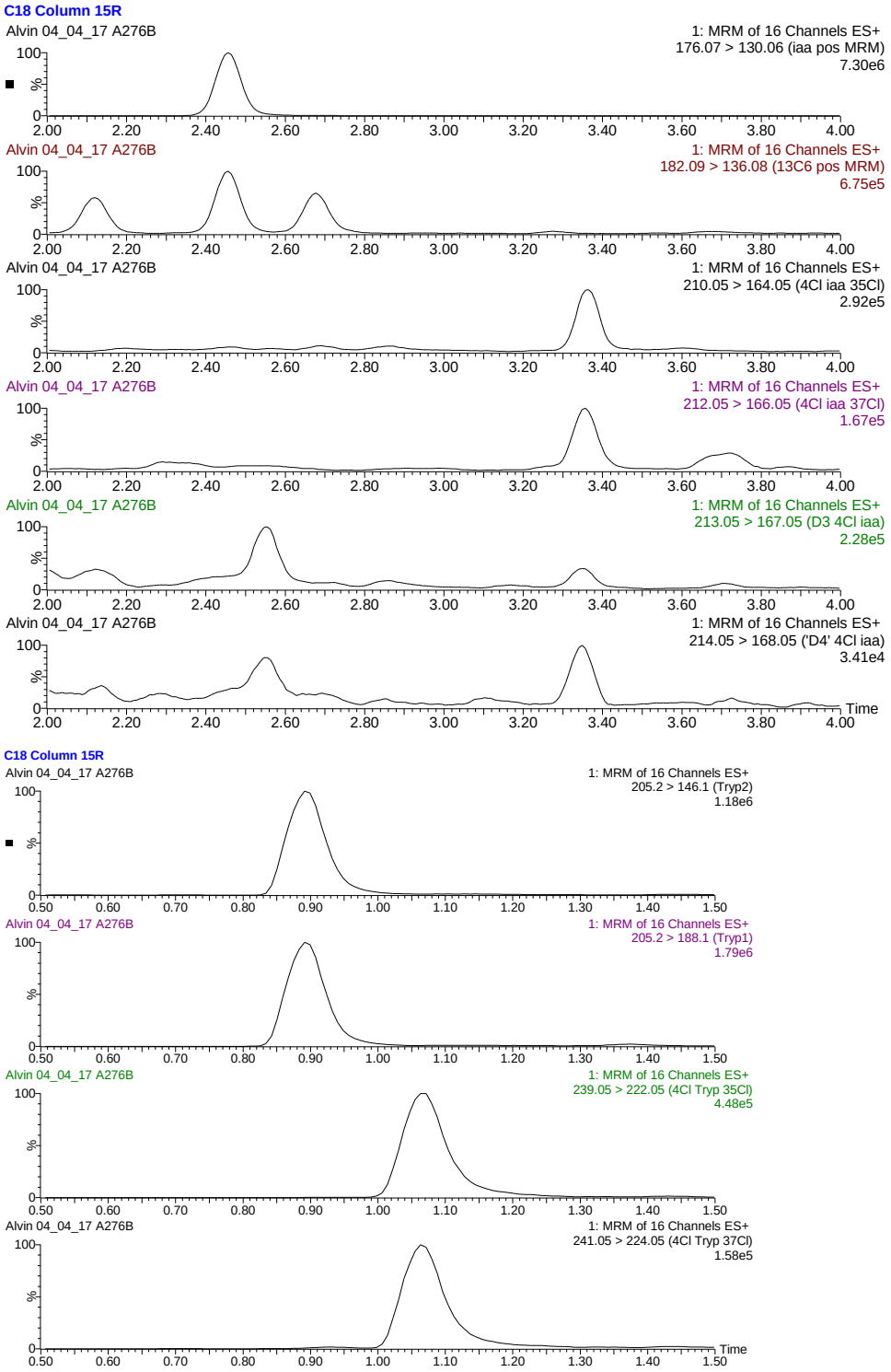


Figure 4.10 UPLC-MS chromatograms (MRM mode) of *Trigonella mesopotamica* (extract of dry seeds) showing the presence of endogenous IAA (2.46 min), 4-Cl-IAA (3.36 min), Trp (0.9 min) and 4-Cl-Trp (1.07 min). Peak intensities (normalised) are plotted against retention time; the height of the strongest peak is shown at the right. Transitions: endogenous IAA, 176-130; [$^{13}C_6$]-IAA, 182-136; endogenous 4-Cl-IAA, 210-164; [$^{2}H_2$]4-Cl-IAA, 212-166; [$^{2}H_3$]4-Cl-IAA, 213-167; [$^{2}H_4$]4-Cl-IAA, 214-168.

Together with the supporting morphological characters (e.g. smooth surface of seed coat, incised stipules margin, notched apex of standard and style longer than ovary), molecular evidence indicates that *Melilotus* is the closest relative of *Trigonella (*Dangi *et al*. 2016). Therefore, it was important to check if loss of chlorinating ability was unique to the genus *Trigonella* or something common to 'chlorinating' genera, just not detected in the past because of a focus on cultivated species. A total of 15 species out of the 16 species of genus *Melilotus* were subjected to 4-Cl-IAA screening. None of the 15 species appeared negative for chlorinated auxin. For instance, the dry seed extract of *Melilotus albus* was found containing 119 ng.g FW^{-1} of total levels of 4-Cl-IAA as well as producing 4-Cl-Trp (Fig. 4.11).

Table 4.2 Endogenous levels of 4-Cl-IAA and IAA in 3 representative species from the genus *Melilotus* (n.d. not detected).

Species	**Tissue type**	**IAA (ng.g^{-1})**	**4-Cl-IAA (ng.g^{-1})**
Melilotus albus	Dry seeds	89	119
Melilotus altissimus	Dry seeds	961	519
Melilotus dentatus	Dry seeds	3001	185
Melilotus elegans	Dry seeds	760	215
Melilotus hirsutus	Dry seeds	2995	287
Melilotus infectus	Dry seeds	542	74
Melilotus italicus	Dry seeds	517	58
Melilotus officinalis	Dry seeds	2467	483
Melilotus segetalis	Dry seeds	707	273
Melilotus siculus	Dry seeds	169	118
Melilotus speciosus	Dry seeds	361	575
Melilotus suaveolens	Dry seeds	995	839
Melilotus sulcatus	Dry seeds	167	130
Melilotus tauricus	Dry seeds	894	283
Melilotus wolgicus	Dry seeds	1395	384

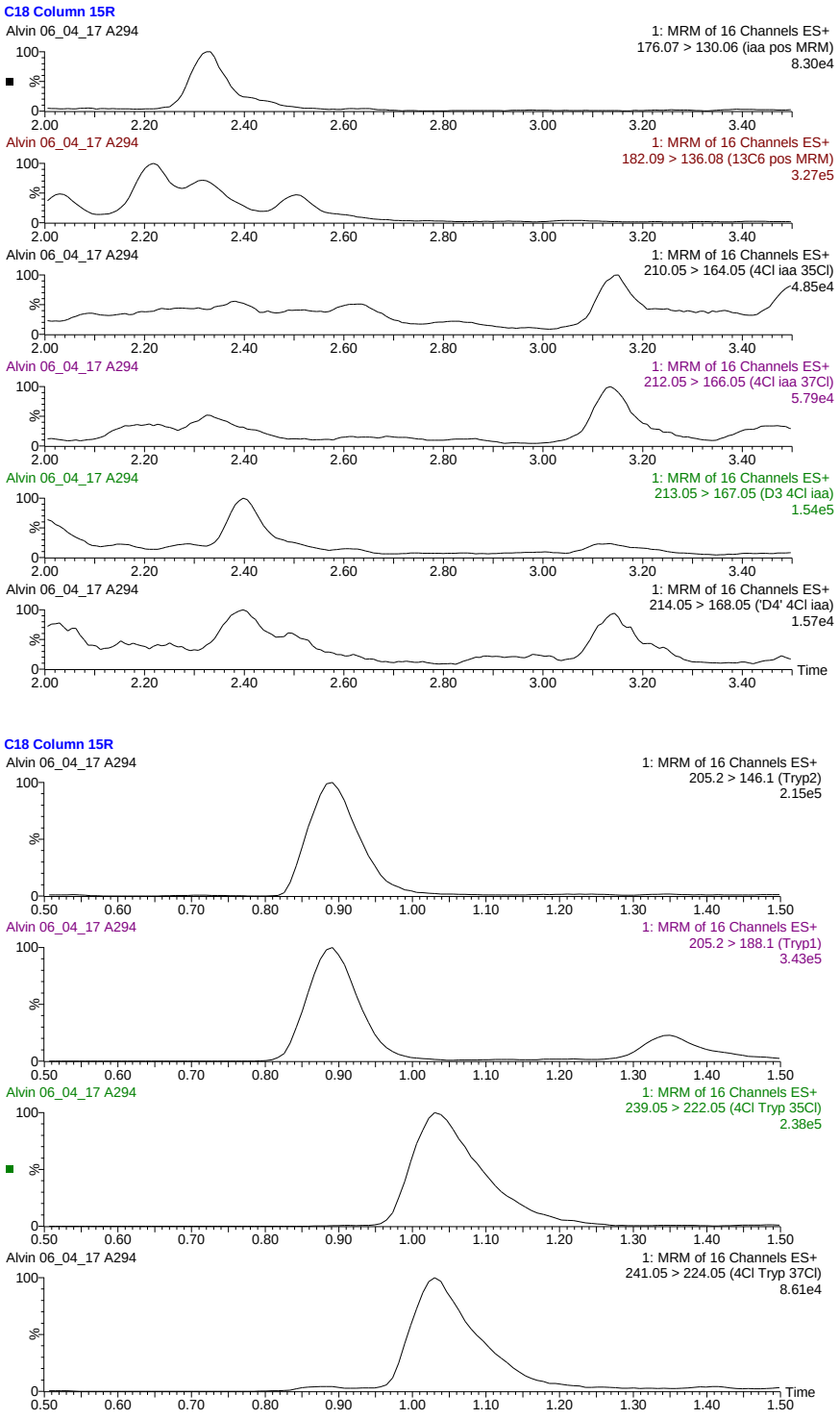


Figure 4.11 UPLC-MS chromatograms (MRM mode) of *Melilotus albus* (extract of dry seeds) showing the presence of endogenous IAA (2.32 min), 4-Cl-IAA (3.14 min), Trp (0.9 min) and 4-Cl-Trp (1.04 min). Peak intensities (normalised) are plotted against retention time; the height of the strongest peak is shown at the right. Transitions: endogenous IAA, 176-130; [$^{13}C_6$]-IAA, 182-136; endogenous 4-Cl-IAA, 210-164; [$^{2}H_2$]4-Cl-IAA, 212-166; [$^{2}H_3$]4-Cl-IAA, 213-167; [$^{2}H_4$]4-Cl-IAA, 214-168.

Taxonomically, after *Melilotus*, *Medicago* is the next most closely related genus to *Trigonella*. Although it has been shown that *Medicago truncatula* produces chlorinated auxin, extensive screening of other representative species from the genus *Medicago* could eliminate the possibility of bias in favour of cultivated species. 7 *Medicago* species were tested for the presence of chlorinated auxin, spanning all major clades within the genus, and all were found to contain both 4-Cl-IAA and 4-Cl- Trp (Table 4.3). For examples, *Medicago noeana* (Fig. 4.12) contains total levels of 4-Cl-IAA as much as 2272 $ng.g(FW)^{-1}$ in dry seeds whereas the dry seed extracts of *M. ruthenica* (Fig. 4.13) and *M. laciniata* (Fig. 4.14) contain 374 and 298 $ng.g(FW)^{-1}$ of total levels of 4-Cl-IAA respectively.

Table 4.3 Endogenous levels of 4-Cl-IAA and IAA, and the presence of 4-Cl-Trp in 7 representative species from the genus *Medicago* (n.d. not detected).

Species	**Tissue type**	**IAA (ng.g^{-1})**	**4-Cl-IAA (ng.g^{-1})**	**4Cl-Trp**
Medicago heyniana	Dry seeds	5	14	Y
Medicago orbicularis	Dry seeds	7	28	Y
Medicago radiata	Dry seeds	13	26	Y
Medicago arabica	Dry seeds	184	60	Y
Medicago laciniata	Dry seeds	173	298	Y
Medicago noeana	Dry seeds	517	2272	Y
Medicago ruthenica	Dry seeds	152	374	Y

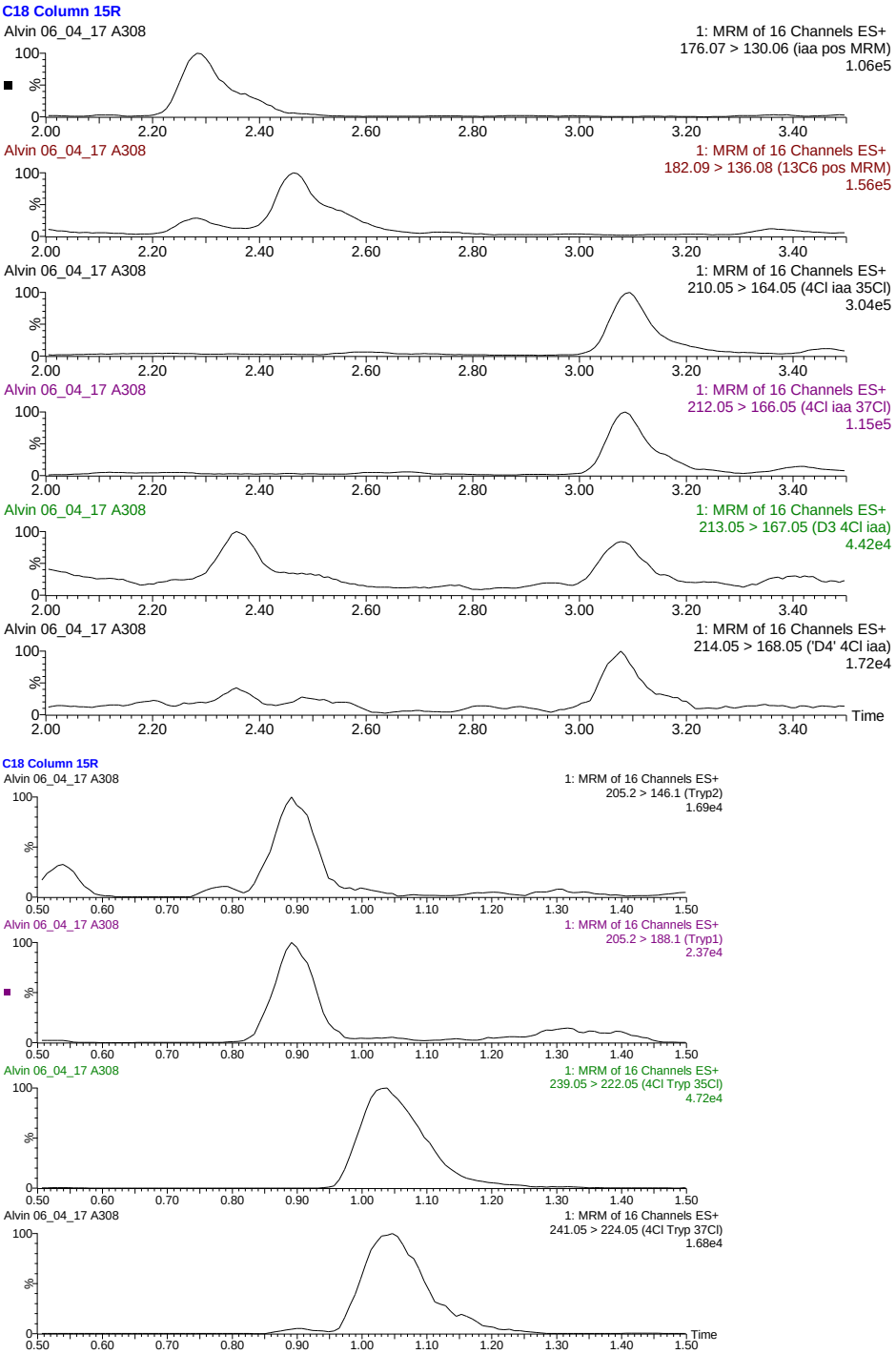


Figure 4.12 UPLC-MS chromatograms (MRM mode) of *Medicago noeana* (extract of dry seeds) showing the presence of endogenous IAA (2.30 min), 4-Cl-IAA (3.10 min), Trp (0.9 min) and 4-Cl-Trp (1.04 min). Peak intensities (normalised) are plotted against retention time; the height of the strongest peak is shown at the right. Transitions: endogenous IAA, 176-130; [$^{13}C_6$]-IAA, 182-136; endogenous 4-Cl-IAA, 210-164; [$^{2}H_2$]4-Cl-IAA, 212-166; [$^{2}H_3$]4-Cl-IAA, 213-167; [$^{2}H_4$]4-Cl-IAA, 214-168.

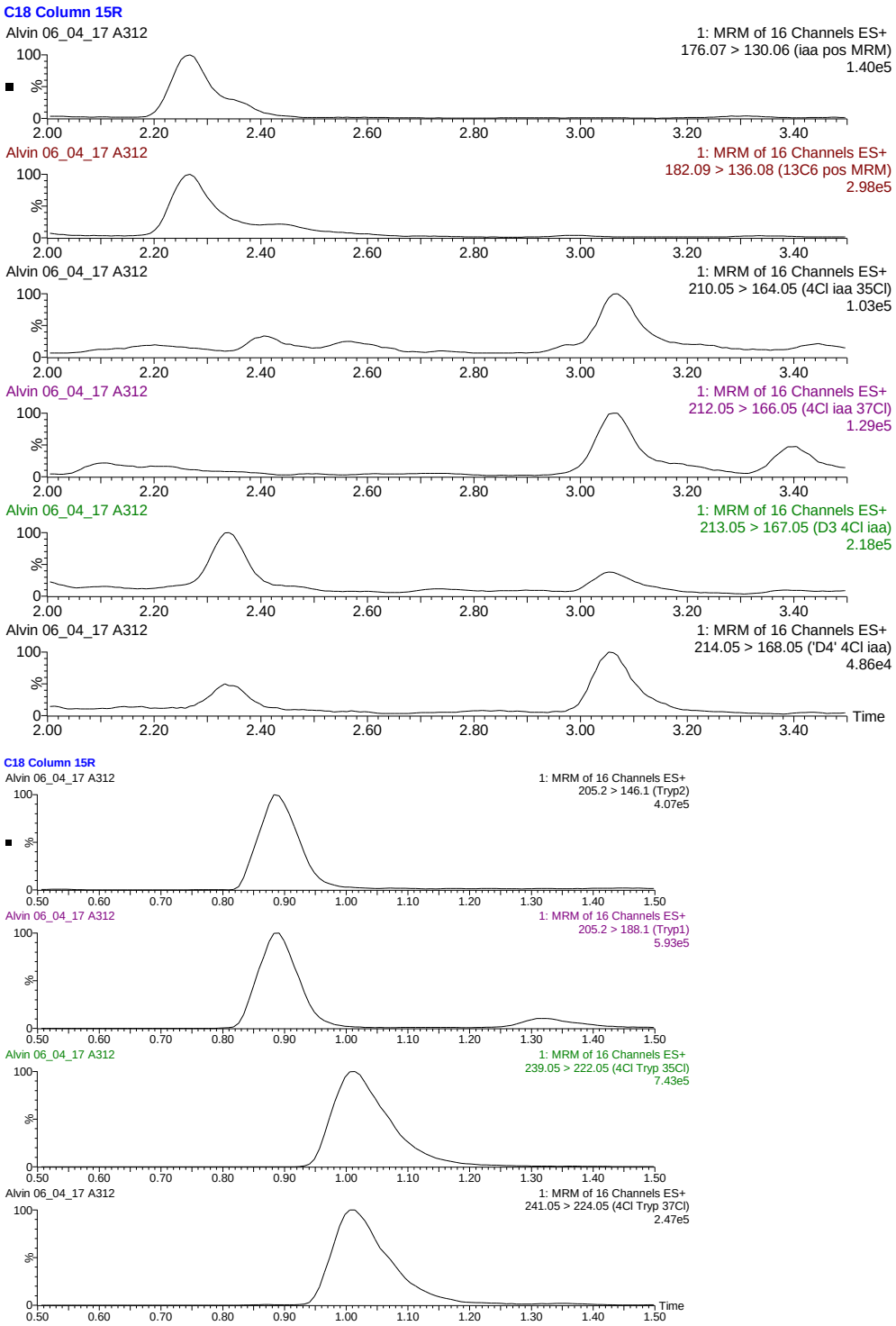


Figure 4.13 UPLC-MS chromatograms (MRM mode) of *Medicago ruthenica* (extract of dry seeds) showing the presence of endogenous IAA (2.26 min), 4-Cl-IAA (3.26 min), Trp (0.88 min) and 4-Cl-Trp (1.01 min). Peak intensities (normalised) are plotted against retention time; the height of the strongest peak is shown at the right. Transitions: endogenous IAA, 176-130; [$^{13}C_6$]-IAA, 182-136; endogenous 4-Cl-IAA, 210-164; [$^{2}H_2$]4-Cl-IAA, 212-166; [$^{2}H_3$]4-Cl-IAA, 213-167; [$^{2}H_4$]4-Cl-IAA, 214-168.

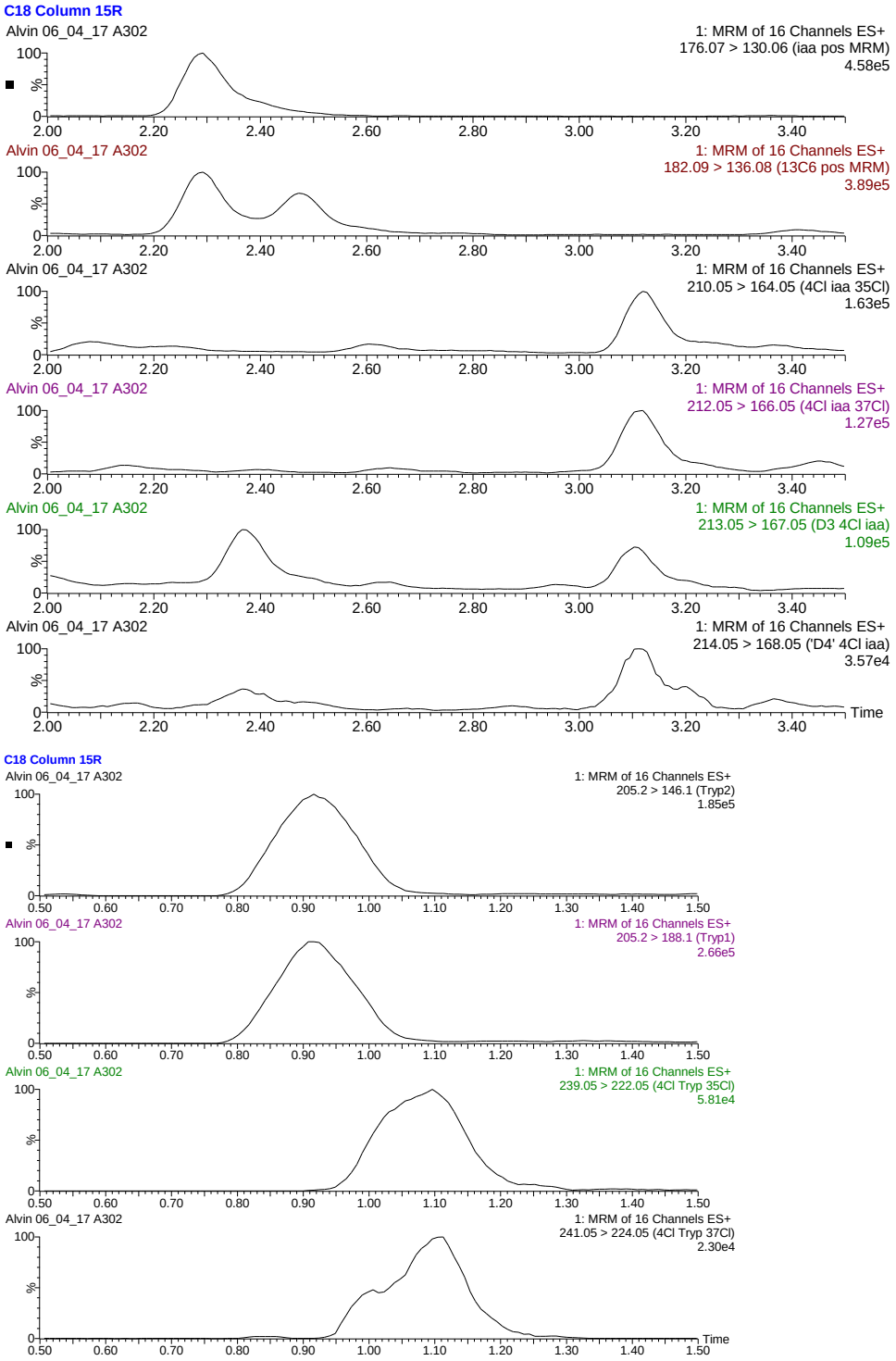


Figure 4.14 UPLC-MS chromatograms (MRM mode) of *Medicago laciniata* (extract of dry seeds) showing the presence of endogenous IAA (2.30 min), 4-Cl-IAA (3.12 min), Trp (0.92 min) and 4-Cl-Trp (1.1 min). Peak intensities (normalised) are plotted against retention time; the height of the strongest peak is shown at the right. Transitions: endogenous IAA, 176-130; [$^{13}C_6$]-IAA, 182-136; endogenous 4-Cl-IAA, 210-164; [$^{2}H_2$]4-Cl-IAA, 212-166; [$^{2}H_3$]4-Cl-IAA, 213-167; [$^{2}H_4$]4-Cl-IAA, 214-168.

In earlier attempts to pin down the evolutionary origin of 4-Cl-IAA within the *Fabeae* and *Trifolieae* (ex. *Parochetus*) tribes, documented in Chapter 3, four species from *Ononis* were found to be producing 4-Cl-IAA as well as 4-Cl-Trp. Here, UPLC-MS/MS showed that another four species of *Ononis* also chlorinate (Table 4.4), including *Ononis alopecuroides* (Fig. 4.15), *O. ornithopodioides* (Fig. 4.16), *O. pubescens* (Fig. 4.17) *amd O. pusilla* (Fig. 4.18).

Table 4.4. Endogenous levels of 4-Cl-IAA and IAA in 4 representative species from the genus *Ononis* (n.d. not detected).

Species	**Tissue type**	**IAA ($ng.g^{-1}$)**	**4-Cl-IAA ($ng.g^{-1}$)**	**4Cl-Trp**
Ononis alopecuroides	Dry seeds	125	441	Y
Ononis ornithopodioides	Dry seeds	19	50	Y
Ononis pubescens	Dry seeds	1092	315	Y
Ononis pusilla	Dry seeds	91	179	Y

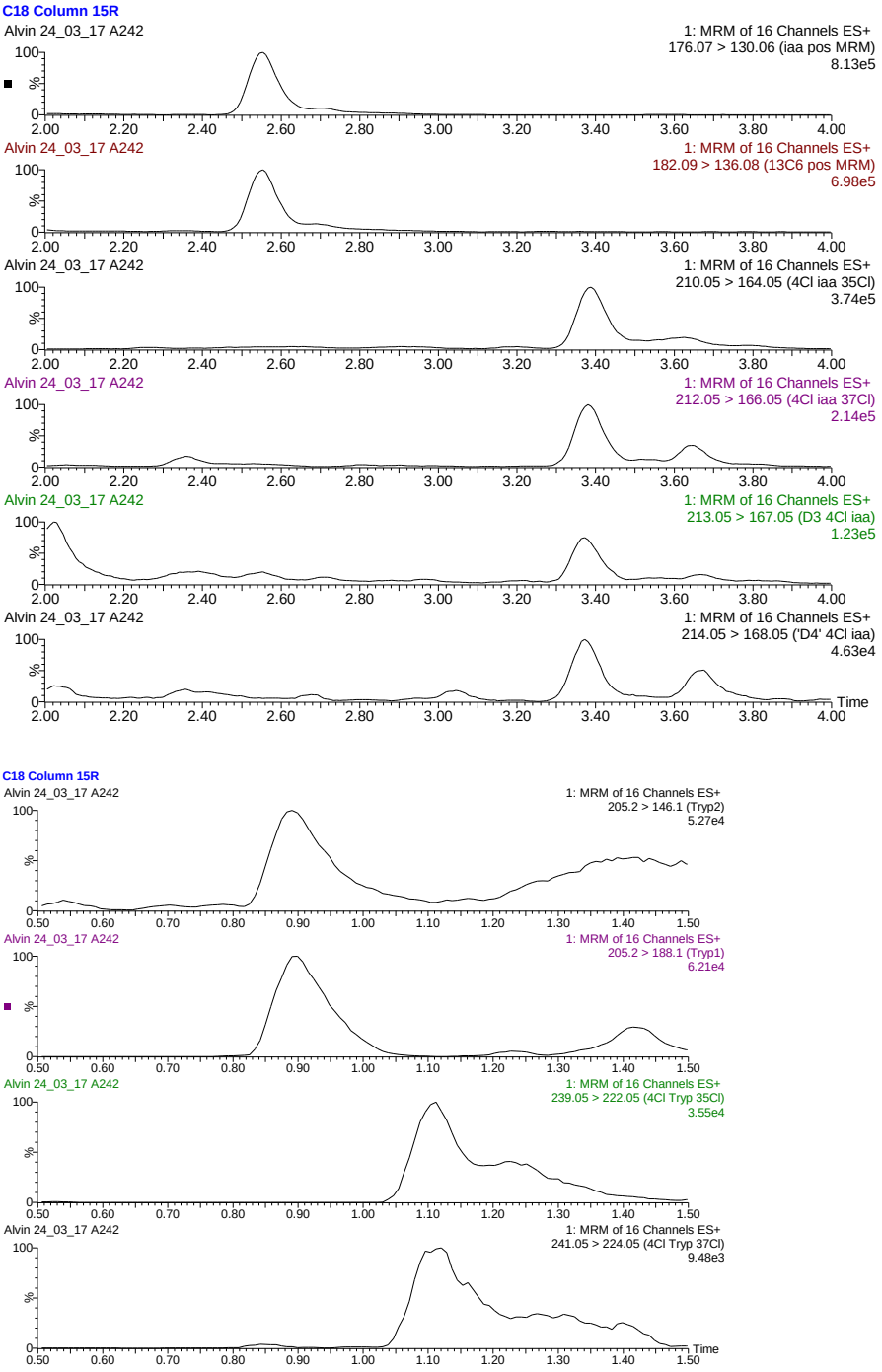


Figure 4.15 UPLC-MS chromatograms (MRM mode) of *Ononis alopecuroides* (extract of dry seeds) showing the presence of endogenous IAA (2.56 min), 4-Cl-IAA (3.38 min), Trp (0.9 min) and 4-Cl-Trp (1.1 min). Peak intensities (normalised) are plotted against retention time; the height of the strongest peak is shown at the right. Transitions: endogenous IAA, 176-130; [$^{13}C_6$]-IAA, 182-136; endogenous 4-Cl-IAA, 210-164; [$^{2}H_2$]4-Cl-IAA, 212-166; [$^{2}H_3$]4-Cl-IAA, 213-167; [$^{2}H_4$]4-Cl-IAA, 214-168.

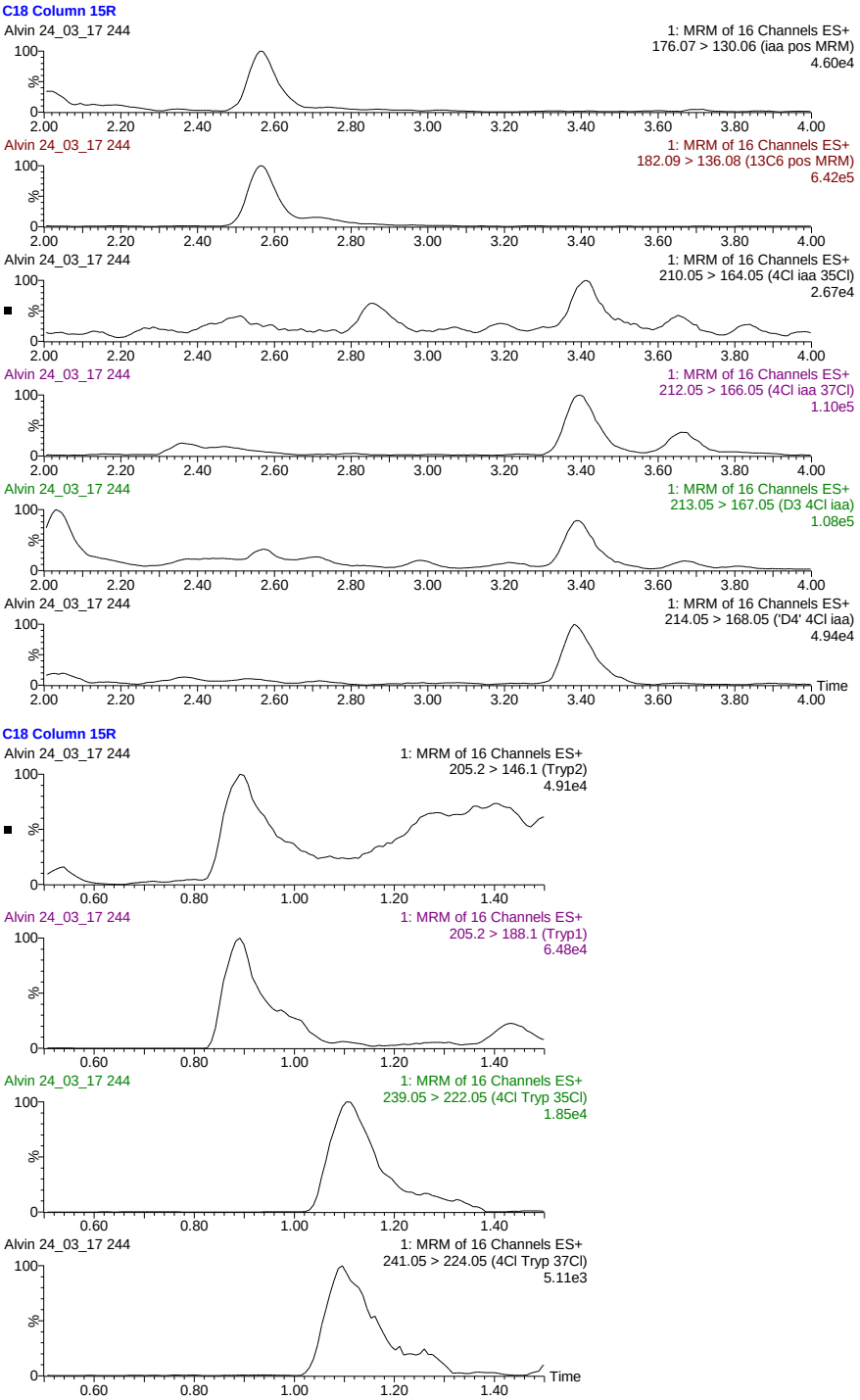


Figure 4.16 UPLC-MS chromatograms (MRM mode) of *Ononis ornithopodioides* (extract of dry seeds) showing the presence of endogenous IAA (2.56 min), 4-Cl-IAA (3.40 min), Trp (0.9 min) and 4-Cl-Trp (1.1 min). Peak intensities (normalised) are plotted against retention time; the height of the strongest peak is shown at the right. Transitions: endogenous IAA, 176-130; [$^{13}C_6$]-IAA, 182-136; endogenous 4-Cl-IAA, 210-164; [$^{2}H_2$]4-Cl-IAA, 212-166; [$^{2}H_3$]4-Cl-IAA, 213-167; [$^{2}H_4$]4-Cl-IAA, 214-168.

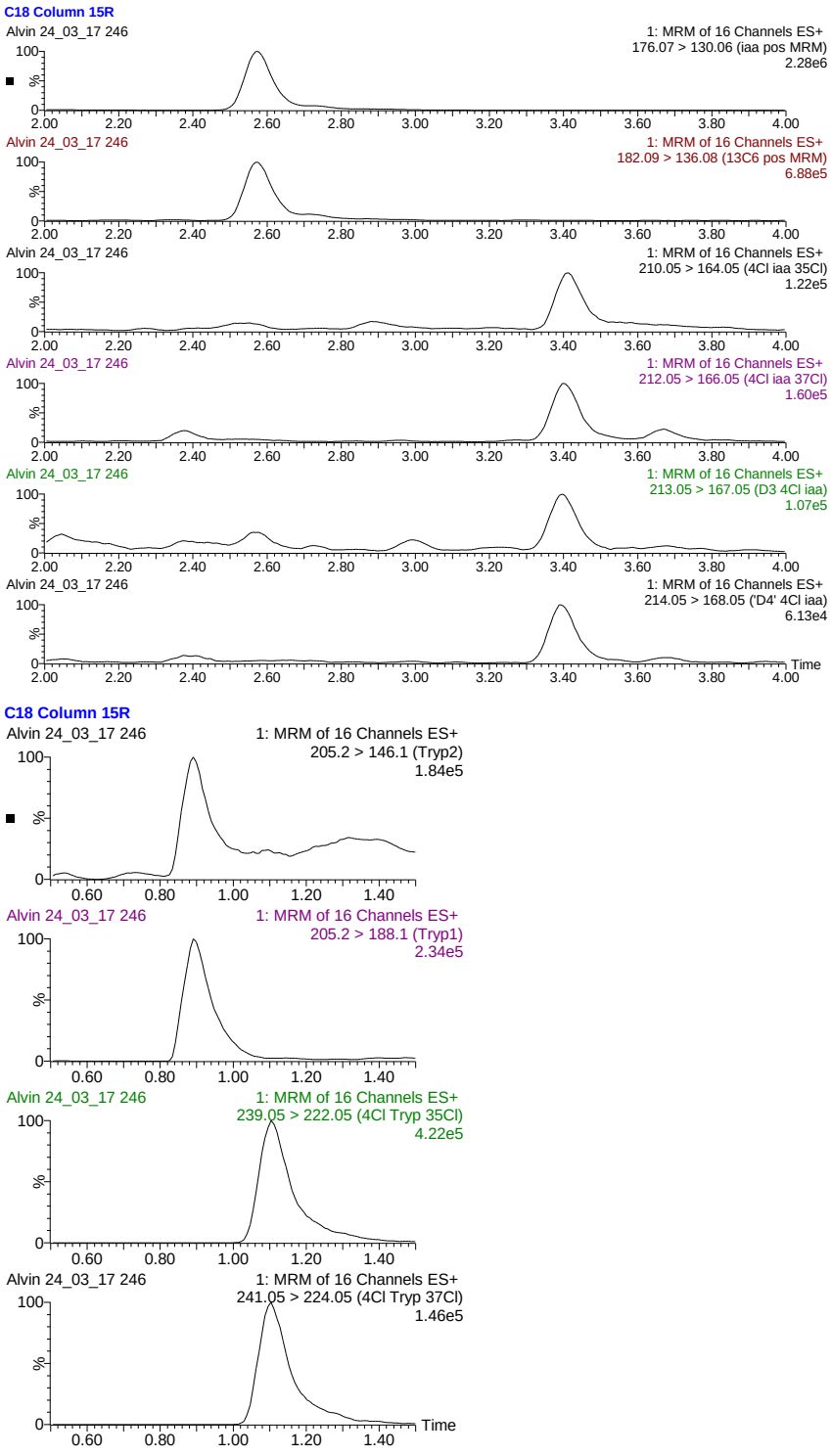


Figure 4.17 UPLC-MS chromatograms (MRM mode) of *Ononis pubescens* (extract of dry seeds) showing the presence of endogenous IAA (2.58 min), 4-Cl-IAA (3.4 min), Trp (0.9 min) and 4-Cl-Trp (1.1 min). Peak intensities (normalised) are plotted against retention time; the height of the strongest peak is shown at the right. Transitions: endogenous IAA, 176-130; [$^{13}C_6$]-IAA, 182-136; endogenous 4-Cl-IAA, 210-164; [$^{2}H_2$]4-Cl-IAA, 212-166; [$^{2}H_3$]4-Cl-IAA, 213-167; [$^{2}H_4$]4-Cl-IAA, 214-168.

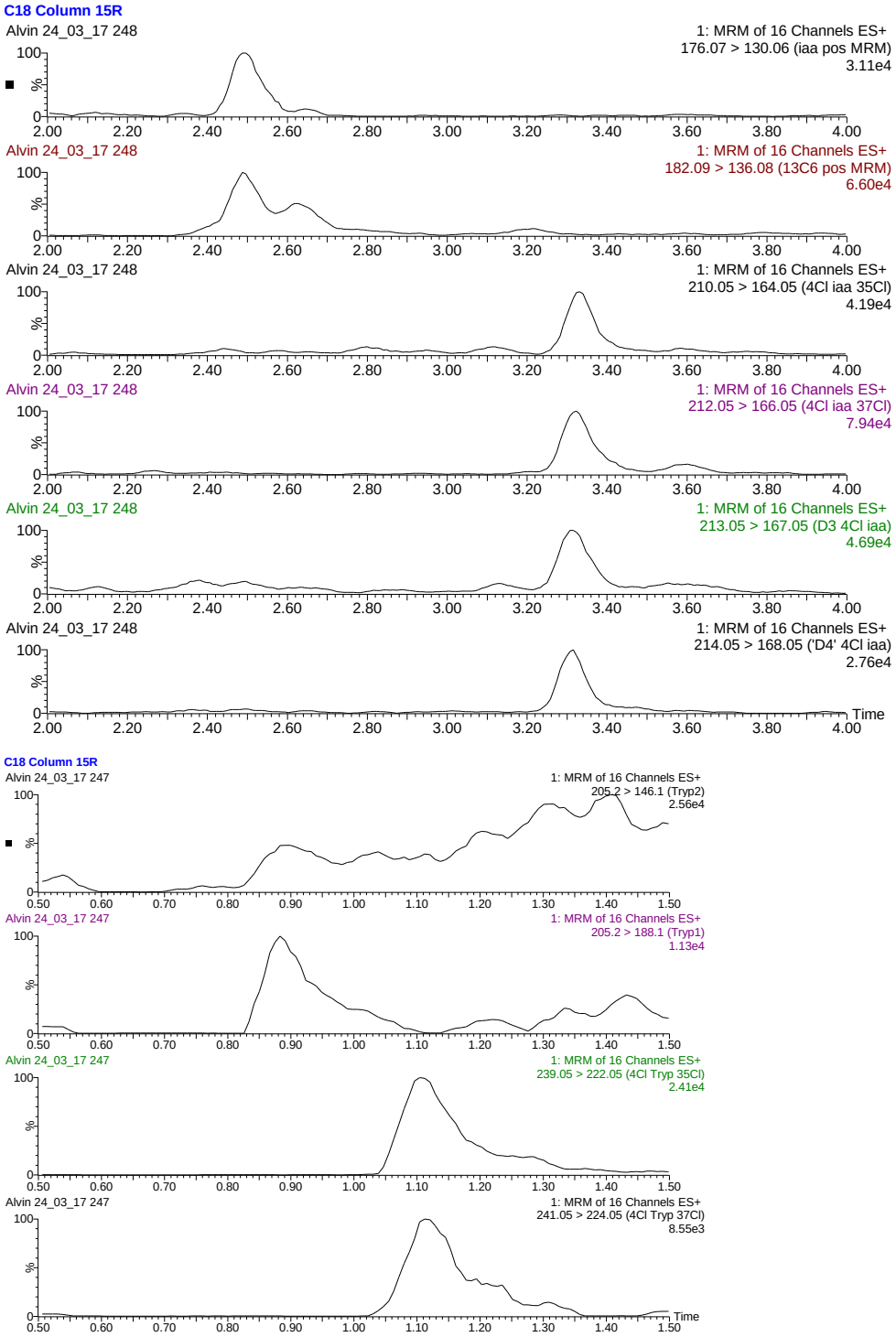
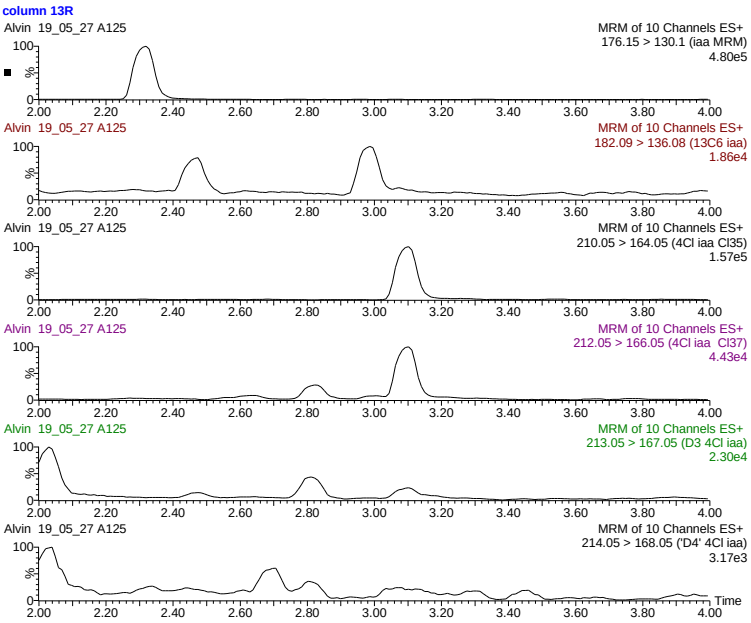


Figure 4.18 UPLC-MS chromatograms (MRM mode) of *Ononis pusilla* (extract of dry seeds) showing the presence of endogenous IAA (2.5 min), 4-Cl-IAA (3.34 min), Trp (0.9 min) and 4-Cl-Trp (1.1 min). Peak intensities (normalised) are plotted against retention time; the height of the strongest peak is shown at the right. Transitions: endogenous IAA, 176-130; [$^{13}C_6$]-IAA, 182-136; endogenous 4-Cl-IAA, 210-164; [$^{2}H_2$]4-Cl-IAA, 212-166; [$^{2}H_3$]4-Cl-IAA, 213-167; [$^{2}H_4$]4-Cl-IAA, 214-168.

In this chapter, the methodology of having separate samples for detection and quantification has proved useful once again especially dealing with such intriguing loss of chlorinating ability in some of the *Trigonella* species but at the same time trying to identify true chlorinating species. For instance, in the absence of [$^{13}C_6$] IAA and [D_2] 4-Cl-IAA internal standards, the corresponding signals of IAA (m/z 176 to 130) as well as the ^{35}Cl-4-Cl-IAA (m/z 210-164) and ^{37}Cl-4-CL-IAA (m/z 212 to 166) did show up indicating the endogenous nature of the compounds detected in the following dry seed extract samples of *T. schlumbergeri* and *T. filipes* (Fig. 4.19 and 4.20).

(A)



(B)

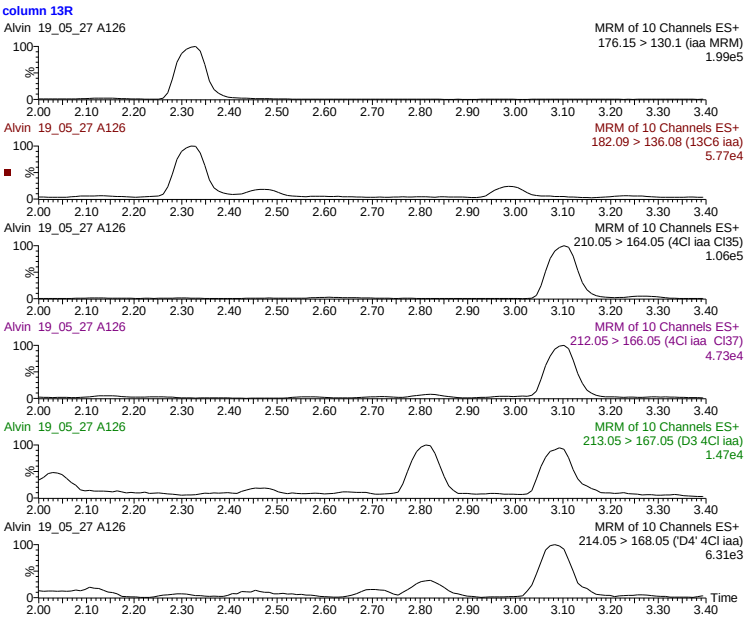
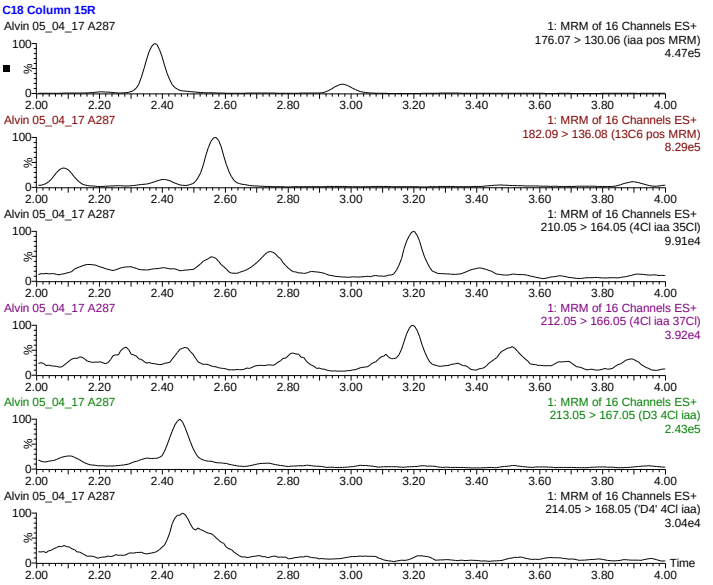


Figure 4.19 UPLC-MS chromatograms (MRM mode) obtained from *Trigonella schlumbergeri* (dry seeds, after hydrolysis), showing separate samples subjected to (A) detection and (B) quantification of the compounds of interest without and with the addition of labelled internal standards ([$^{13}C_6$]-IAA and [D_2]-4-Cl-IAA). RTs for the compounds of interests are as follows: IAA = [$^{13}C_6$]-IAA = 2.32 min, 4-Cl-IAA = [D_2]-4-Cl-IAA = 3.1 min. Peak intensities (normalised) are plotted against retention time; the height of the strongest peak is shown at the right. Transitions: endogenous IAA, 176-130; [$^{13}C_6$]-IAA, 182-136; endogenous 4-Cl-IAA, 210-164; [$^{2}H_2$]4-Cl-IAA, 212-166; [$^{2}H_3$]4-Cl-IAA, 213-167; [$^{2}H_4$]4-Cl-IAA, 214-168.

(A)



(B)

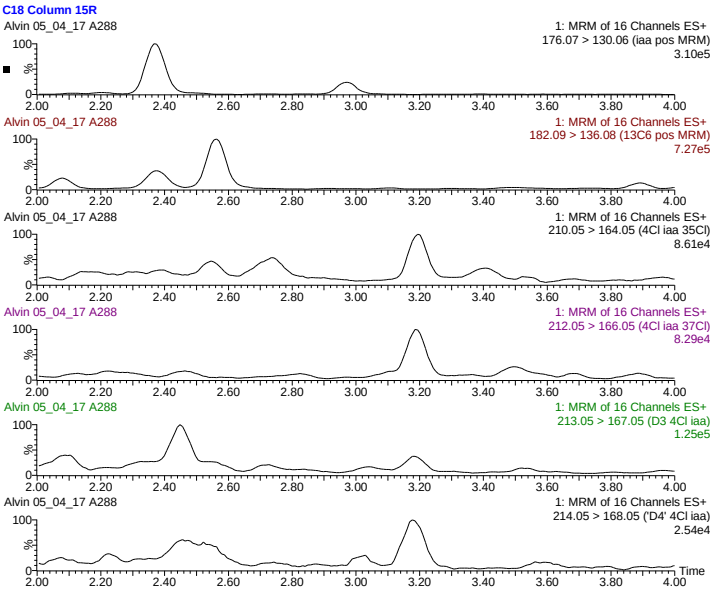


Figure 4.20 UPLC-MS chromatograms (MRM mode) obtained from *Trigonella filipes* (dry seeds, after hydrolysis), showing separate samples subjected to (A) detection and (B) quantification of the compounds of interest, without and with the addition of labelled internal standards ([$^{13}C_6$]-IAA and [D_2]-4-Cl-IAA). RTs for the compounds of interests are as follows: IAA = [$^{13}C_6$]-IAA = 2.38 min, 4-Cl-IAA = [D_2]-4-Cl-IAA = 3.2 min. Peak intensities (normalised) are plotted against retention time; the height of the strongest peak is shown at the right. Transitions: endogenous IAA, 176-130; [$^{13}C_6$]-IAA, 182-136; endogenous 4-Cl-IAA, 210-164; [$^{2}H_2$]4-Cl-IAA, 212-166; [$^{2}H_3$]4-Cl-IAA, 213-167; [$^{2}H_4$]4-Cl-IAA, 214-168.

Three species of *Trifolium* were also subjected to the screening for 4-Cl-IAA and 4-Cl-Trp, namely *Trifolium alexandrium*, *T. glomeratum* and *T. resupinatum*. Chlorinated auxin and tryptophan were detected in all three tested species (Figs. 4.21, 4.22 and 4.23). Together with the previously studied species (Chapter 2), all six *Trifolium* tested so far are, indeed, chlorinating species.

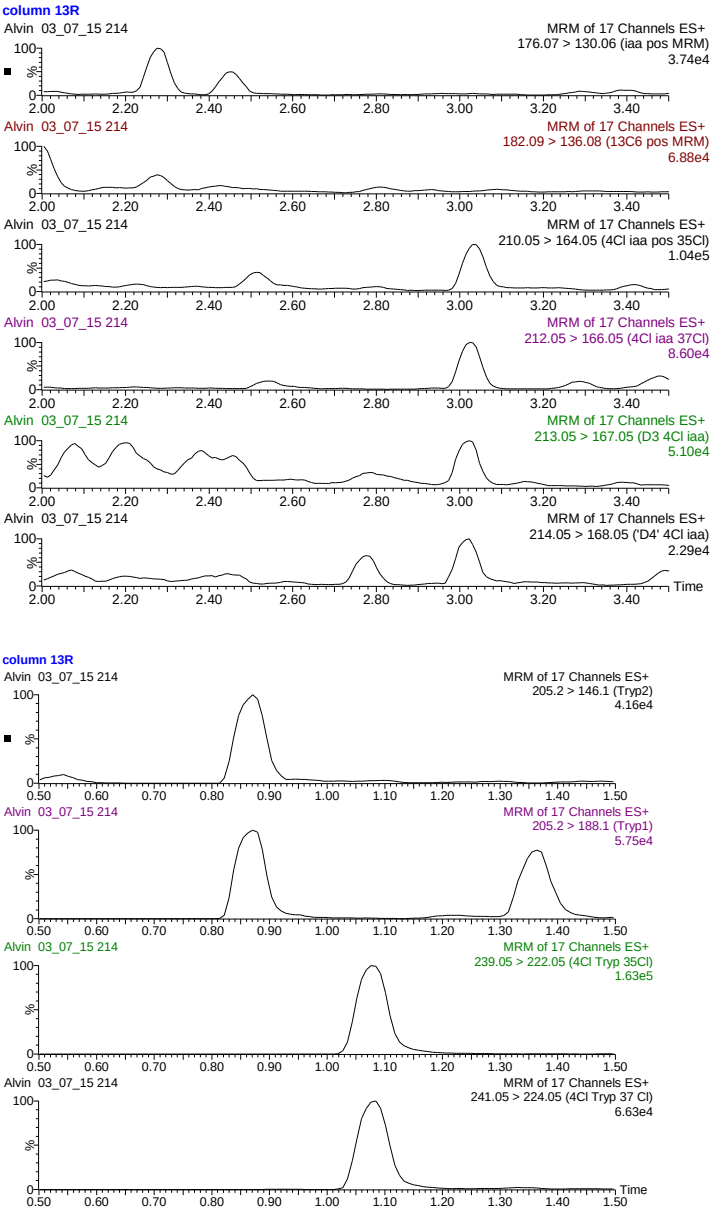


Figure 4.21 UPLC-MS chromatograms (MRM mode) of *Trifolium alexandrium* (extract of dry seeds) showing the presence of endogenous IAA (2.28 min), 4-Cl-IAA (3.2 min), Trp (0.86 min) and 4-Cl-Trp (1.06 min). Peak intensities (normalised) are plotted against retention time; the height of the strongest peak is shown at the right. Transitions: endogenous IAA, 176-130; [$^{13}C_6$]-IAA, 182-136; endogenous 4-Cl-IAA, 210-164; [$^{2}H_2$]4-Cl-IAA, 212-166; [$^{2}H_3$]4-Cl-IAA, 213-167; [$^{2}H_4$]4-Cl-IAA, 214-168.

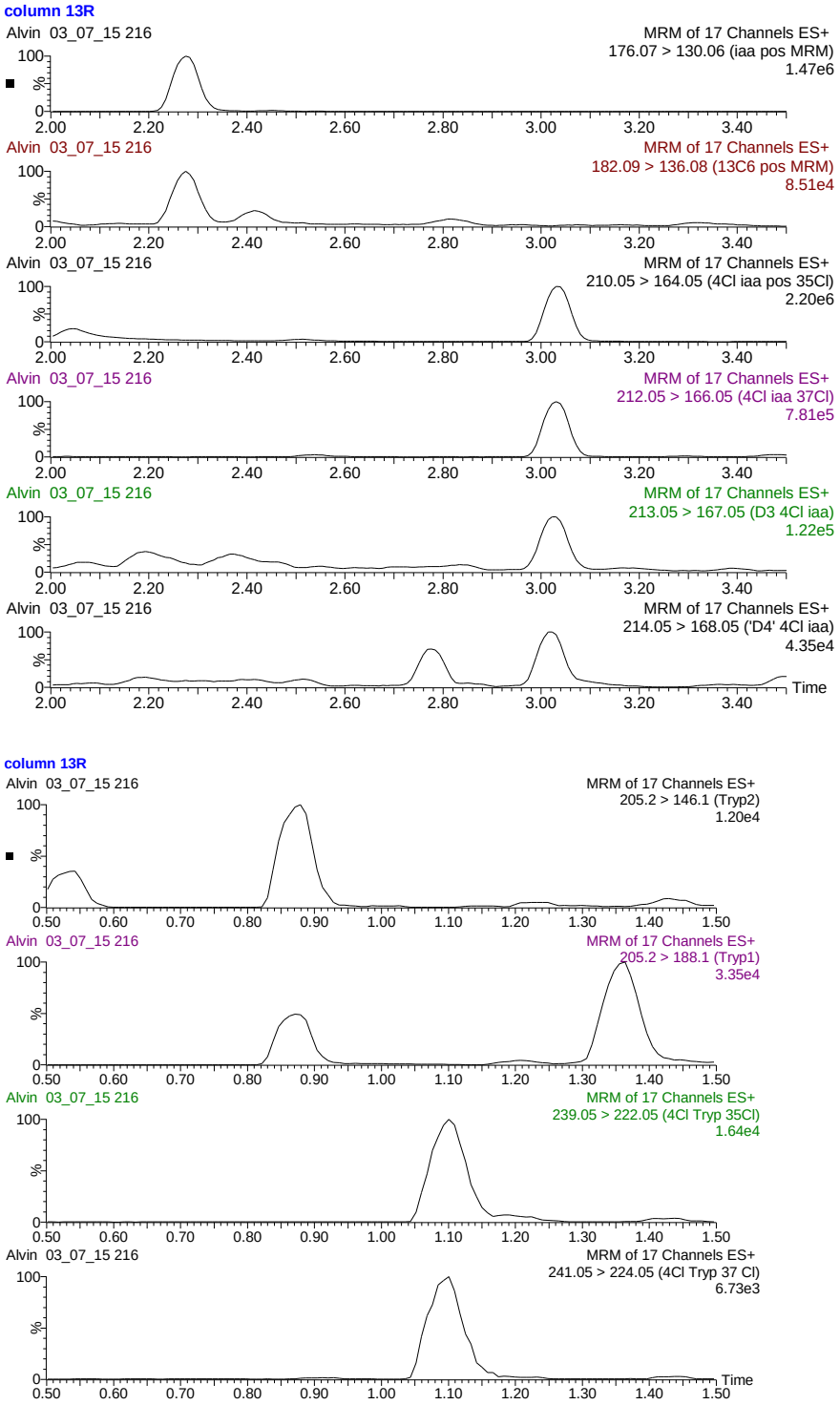


Figure 4.22 UPLC-MS chromatograms (MRM mode) of *Trifolium glomeratum* (extract of dry seeds) showing the presence of endogenous IAA (2.28 min), 4-Cl-IAA (3.04 min), Trp (0.88 min) and 4-Cl-Trp (1.1 min). Peak intensities (normalised) are plotted against retention time; the height of the strongest peak is shown at the right. Transitions: endogenous IAA, 176-130; [$^{13}C_6$]-IAA, 182-136; endogenous 4-Cl-IAA, 210-164; [$^{2}H_2$]4-Cl-IAA, 212-166; [$^{2}H_3$]4-Cl-IAA, 213-167; [$^{2}H_4$]4-Cl-IAA, 214-168.

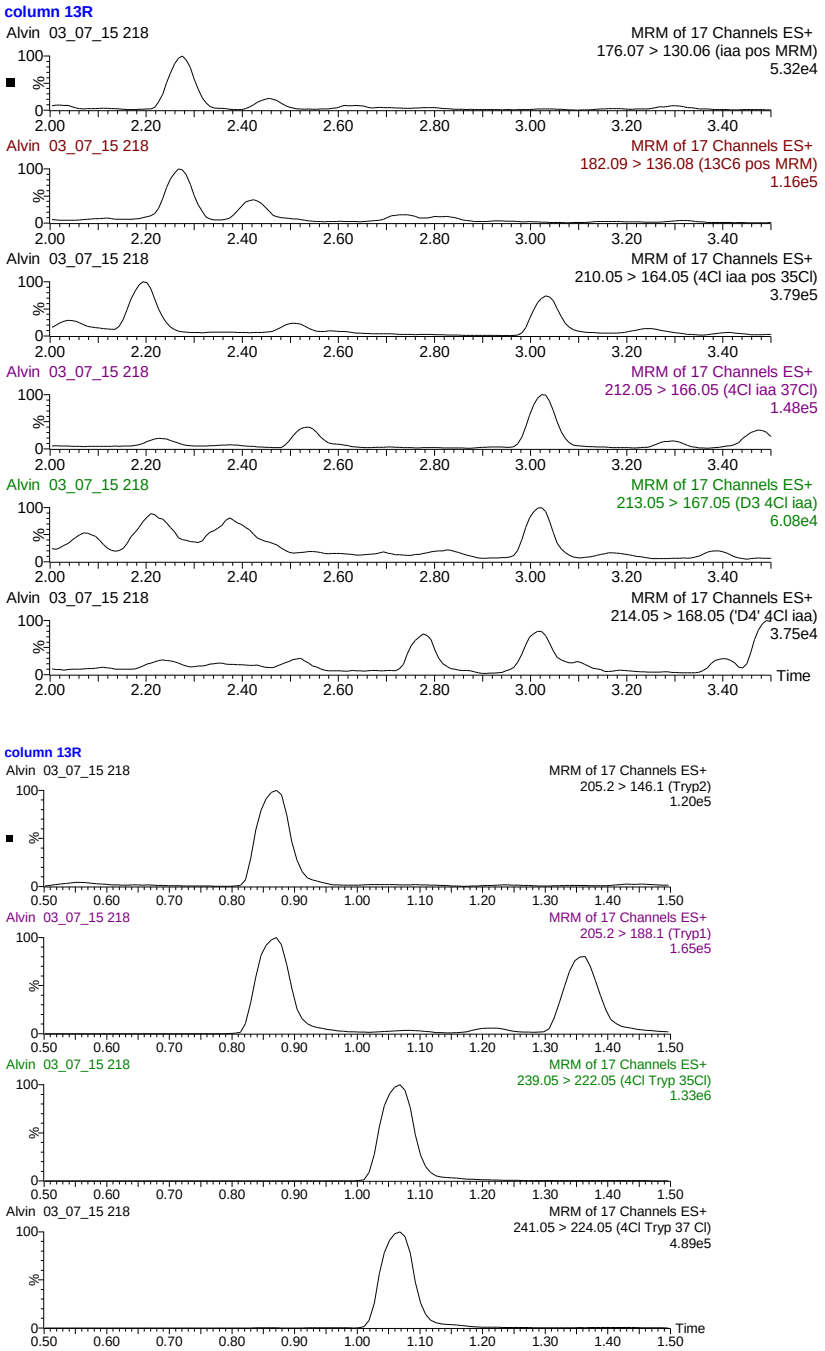


Figure 4.23 UPLC-MS chromatograms (MRM mode) of *Trifolium resupinatum* (extract of dry seeds) showing the presence of endogenous IAA (2.27 min), 4-Cl-IAA (3.04 min), Trp (0.87 min) and 4-Cl-Trp (1.07 min). Peak intensities (normalised) are plotted against retention time; the height of the strongest peak is shown at the right. Transitions: endogenous IAA, 176-130; [$^{13}C_6$]-IAA, 182-136; endogenous 4-Cl-IAA, 210-164; [$^{2}H_2$]4-Cl-IAA, 212-166; [$^{2}H_3$]4-Cl-IAA, 213-167; [$^{2}H_4$]4-Cl-IAA, 214-168.

This comprehensive screening approach spanning the closely related genera *Melilotus*, *Trifolium*, *Medicago* and *Ononis* suggests that the presence of some non-chlorinating species in the genus *Trigonella* does not seem to be universal among genera in the *Fabeae-Trifolieae*. The homeoplastic non-chlorinating trait among the *Trigonella* species is probably due to two losses of function in parallel within the genus in terms of tryptophan chlorinating ability. *Trigonella* is a species rich genus that loss or gain of chlorinating ability may provide adaptive flexibility. Although it is unclear why some species only chlorinate tryptophan but not further produce 4-Cl-IAA, such chlorinating and non-chlorinating traits may have an effect on seed size, pod architecture and starch contents in the *Trigonellas.*

Chapter 5 Mutant-based Study of a Vanadium-dependent Haloperoxidase Gene in *Pisum sativum*

5.1 INTRODUCTION

Despite exhibiting similar biological functions but distinct structures, six families of halogenating enzymes, which might have evolved independently, have been identified thus far: cofactor-free haloperoxidases (HPO), vanadium-dependent haloperoxidases (VHPO), heme iron-dependent halogenases (F-HG), non-heme iron-dependent halogenases (NI-HG), flavin-dependent halogenases (F-HG), and S-adenosyl-L-methionine (SAM)-dependent halogenases (S-HG) (Xu and Wang, 2016). Based on phylogenetic and structural analyses, these HPO, VHPO, HI-PO, NI-HG, F-HG, and S-HG halogenating enzyme families are believed to have evolutionary relationships to the α/β hydrolases, acid phosphatases, peroxidases, chemotaxis phosphatases, oxidoreductases, and SAM hydroxide adenosyltransferases, respectively (Xu and Wang, 2016).

Although a plethora of tryptophan-derived metabolites are produced by plants, including 4-Cl-IAA, naturally occurring halogenated compounds are actually rare in planta. In fact, the repertoire of tryptophan halogenases has well been characterised in bacteria, cyanobacteria, seaweeds and fungi, yet not a single study has discovered any endogenous chlorinating enzymes in plants to date. Furthermore, among all studied tryptophan halogenases, the regioselective chlorination of tryptophan occurs in the 7- (PrnA from *Pseudomonas fluorescens*; RebH from *Lechevalieria aerocolonigenes*), 6- (ThdH or Thal from *Streptomyces albogriseolus*), 5- (PyrH from *Streptomyces rugosporus*) and 2-positions (CmdE from the myxobacterium Chondromyces crocatus) of the indole ring (Wagner *et al.*, 2009; Weichold *et al.*, 2016) (Fig. 5.1) (Table 5.1), but not in the 4-position.

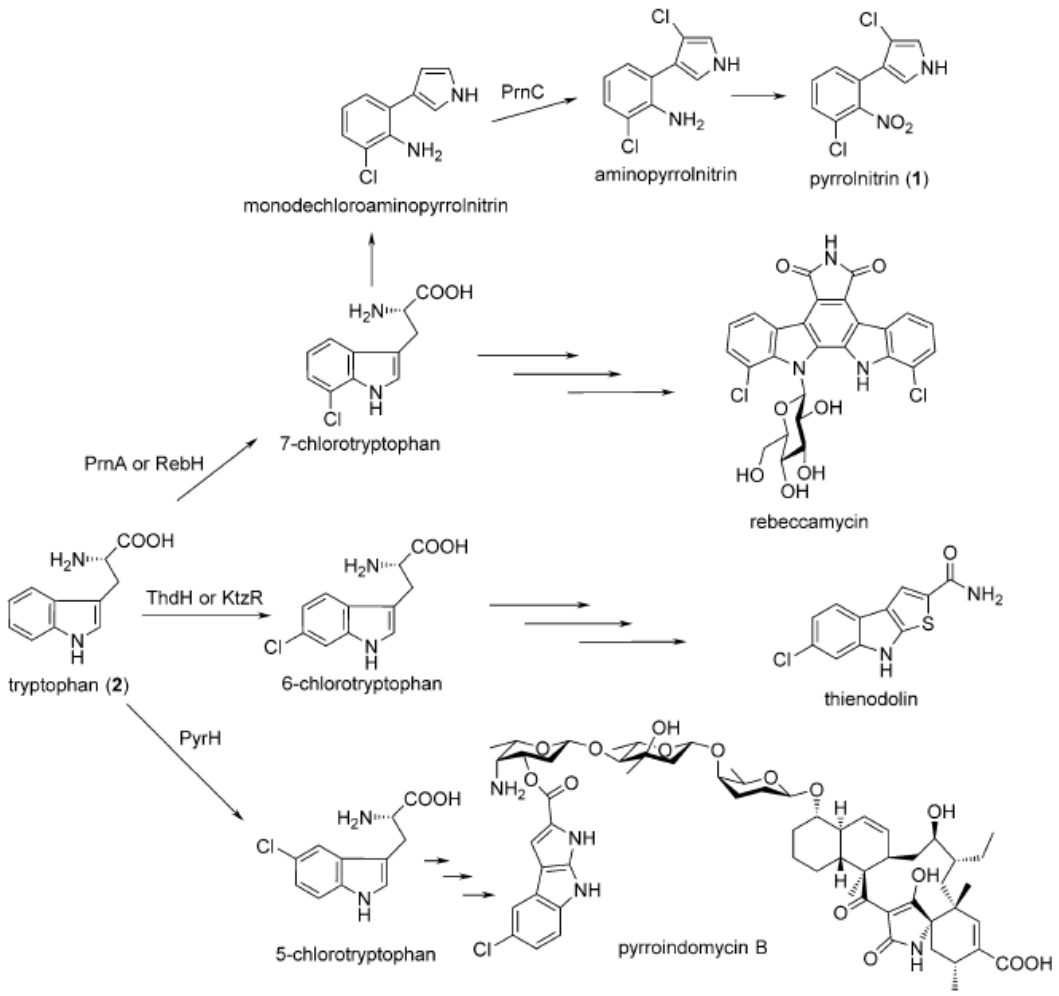


Figure 5.1 Chlorination of tryptophan by flavin-dependent tryptophan halogenases with different regioselectivities during the biosysthesis of pyrrolnitrin, rebeccamycin, thienodolin and pyrroindomycin. From Weichold *et al*. 2016.

Table 5.1 Flavin-dependent halogenases that halogenate tryptophan

Halogenase	Natural product	Biological organism	Halogenated structural element	GenBank accession number
PrnA	pyrrolnitrin	*Pseudomonas fluorescens* BL915	Tryptophan at position C-7	PFU74493
RebH	rebeccamycin	*Lechevalieria aerocolonigenes* ATCC 39243	Tryptophan at position C-7	AJ414559
Thal	thienodolin	*Streptomyces albogriseolus*	Tryptophan at position C-6	EF095207
PyrH	pyrroindomycin B	*Streptomyces rugosporus* LL-42D005	Tryptophan at position C-5	AY623051
CmdE	chondramides B/D	*Chondromyces crocatus* Cm c5	Tryptophan at position C-2	AM179409

The chlorinating ability of specific genera in the Fabaceae promises an exciting research area that will eventually elucidate the biosynthesis of chlorinated auxin, particularly the chlorination step and the chlorinating enzyme(s) involved. Further investigations are needed to find what is or are the halogenating enzymes in the Fabaceae, chlorinating the 4-position of the indole ring. Given the potency of halogenated compounds, understanding 4-Cl-IAA biosynthesis may unlock further potential of grain legumes and pasture legumes that are economically important to feed the hungry world.

Recently Alves-Carvalho *et al.* (2015) utilised high-throughput next-generation sequencing to make an extensive de novo assembly of RNA-seq data from the pea cultivar 'Cameor'. This approach discovered almost all expressed genes in pea plant tissues, by sequencing normalised libraries for seeds, meristems, flowers, cotyledons, seedlings, epicotyls, hypocotyls, immature and mature leaves, stems, immature pods and seeds, and young plantlets, as well as below ground plant tissues (Alves-Carvalho *et al.*, 2015). A Unigene set of expressed sequences at full length was produced after de novo assembly of high-throughput sequences obtained for the various tissue types mentioned, and for a range of developmental stages and nutritive conditions (Alves-Carvalho *et al.*, 2015). Then, the pea RNA-seq gene atlas was developed using that Unigene set. The deployment of the pea gene atlas portal allows us to search for annotated, halogenase-like candidate gene(s). For instance, 16 acid phosphatase/vanadium-dependent haloperoxidase-related peptide sequences were found when a search was conducted within the online pea RNA-seq gene atlas database.

Previously, Targeting Induced Local Lesions in Genomes, or TILLING, as described by Dalmais *et al.* (2008), successfully generated *PsTAR2* mutants by isolating a *tar2* knockout mutant from an ethyl methanesulfonate mutant collection of 4,800 pea lines on the genetic background of 'Cameor'. The effectiveness of this approach has been proven and a full characterisation of the *tar2* mutation has been published (McAdam *et al.*, 2017), documenting the levels of IAA and 4-Cl-IAA at different seed developmental stages. Thus, TILLING and the created mutants facilitated the investigation of the role of pea aminotransferases in 4-Cl-IAA biosynthesis *in vivo*, complementing the *in vitro* analyses of PsTAR1 and PsTAR2 in *Pisum sativum*.

TILLING all the 16 candidate genes annotated as vanadium-dependent haloperoxidase by Alves- Carvalho *et al*. (2015) was not feasible. Instead, one candidate VHPO sequence was chosen, based on phylogenetic relationship of putative VHPO gene sequences from pea, *Medicago*, *Lotus*, *Cicer*, *Trifolium* and other plants. In the TILLING screen, mutation(s) affecting the chosen VHPO gene and the nature of the mutation(s) will be identified by sequencing. Eventually, two lines of mutant were obtained by TILLING, conducted by Dr. Marion Dalmais, Dr Abdelhafid Bendahmane and their team at the Unité de Recherche en Génomique Végétale, Evry, France. In this study, a VHPO-like candidate gene and its mutants were investigated to check if it is a tryptophan/auxin halogenase in the Fabaceae.

5.2 Materials and methods

5.2.1 Sequence availability

Three pea VHPO gene sequences (VHPO1, VHPO2 and VHPO3) were obtained from the 'Caméor' Unigene set consisting of annotated pea orthologs available online through The Pea RNA-Seq gene atlas (http://bios.dijon.inra.fr/FATAL/cgi/pscam.cgi) (Alves-Carvalho *et al.* 2015). Details of sequences shown in Table 5.2 were used in this study for phylogenetic analysis.

Table 5.2 Details of vHPO related sequences of various plants used for phylogenetic tree construction

Species	Gene name	Sequence
Pisum sativum	PsVHPO1	PsCam021226
	PsVHPO2	PsCam043648
	PsVHPO3	PsCam021353
Medicago truncatula	MtVHPO1	XM_003601766
	MtVHPO2	XM_003596891
	MtVHPO3	TC202187
	Mt_1	XM_003596891
	Mt_2	BT134107
	Mt_out	XM_003595828
Lotus japonicus	LjVHPO1	BT135522
	LjVHPO2	CM0105.840
	LjVHPO3	CM0105.840
	Lj_1	BT140406
	Lj_out	CM0996.190
Cicer arietinum	Ca_1	XM_004502262
	Ca_2	XM_004502261
	Ca_3	XM_004487375
	Ca_4	XM_004490438
	Ca_5	XM_004497416
Trifolium pratense	Tp_1	GAOU01044928
Theobroma cacao	Tc_1	XM_007051484

5.2.2 Plant materials

Wild-type (*VHPO1*) pea plants used for this study was the dwarf line Caméor. Mutant (*vhpo1*) pea plants originated from TILLING program in France were also in the genetic background of the dwarf line Caméor. All plants were grown as described previously (Jager *et al.*, 2007). Seeds from these plants were harvested for the quantification of IAA and 4-Cl-IAA levels.

5.2.3 Primers

Degenerate primers for gene isolation were designed using CODEHOP strategy (Rose *et al.*, 1998) on conserved domains identified from protein sequence alignments using BLOCK MAKER application (http://blocks.fhcrc.org/codehop.html). Selected primers were manually optimised by reference to the input legume sequences. Other primers were designed using Primer3 (http://bioinfo.ut.ee/primer3-0.4.0/) based on legume sequences to target conserved regions in pea. The details of primer sequences are recorded in Appendix 1.

5.2.4 Gene isolation

5.2.4.1 DNA extraction

Genomic DNA was isolated from individual plants using a modified CTAB protocol. Harvested tissue was frozen in liquid nitrogen and ground into fine powder in a mill mixer by adding tungsten carbide bead to each tube. 500µL of extraction buffer (100mM Tris-HCl pH8, 1.4M NaCl, 20mM EDTA, 2% w/v CTAB, 20mM 2-β-mercaptoethanol) was added into each tube prior to incubation at 60°C for 15 min. 500µL of Chloroform:Isoamyl alcohol mix (24:1) was then added to the tubes, and the contents were mixed by gentle inversion. The samples were centrifuged at 14000 rpm for 1 min and the upper aqueous phase was transferred to new tubes and again extracted. 1mL of precipitation buffer (50mM Tris-HCL pH8, 10mM EDTA, 1%w/v CTAB) was added after the second extraction. The contents of the tube were mixed gently and incubated at room temperature for 10-15 min to allow the formation of thread-like precipitate. The precipitate was collected by centrifugation at 14000 rpm for 10min and dissolved in 300µL of 1.5M NaCl containing 1µL RNase A (25mg/mL). The solution was incubated at 50°C for 15 min or until the pellet was fully dissolved. DNA was precipitated by adding 600µL of 100% ethanol and

collected by centrifugation at 14000 rpm for 10 min. After centrifugation, the DNA pellet was washed in 200µL of 70% ethanol, air-dried and resuspended in 50µL of sterilised water.

5.2.4.2 DNA quantification

DNA concentrations were determined using a NanoDrop 8000 Spectrophotometer (Thermo Scientific, https://www.thermofisher.com/order/catalog/product/ND-8000-GL)

5.2.4.3 Polymerase Chain Reaction (PCR)

Most PCR reactions were performed in 50µL, which included 5µL of DNA template and 45µL of master mix. The master mix was prepared according to Table 5.3 and scaled up to the required number of reactions.

Table 5.3 Reagents required for a 50µL PCR reaction.

Reagent	For 1 reaction (50µL)
DNA template (added separately)	5µL
PCR buffer (5x)	10µL
dNTPs (10mM)	1µL
Forward primer (10mM)	1µL
Reverse primer (10mM)	1µL
DNA polymerase	0.2µL
Sterilised water	31.8µL

PCR was carried out in a thermal cycler with heated lid as follows: an initial template denaturation step at 95°C for 1 min, followed by 35 cycles of 95°C for 15seconds (denaturation of DNA into single strands), T_m (annealing temperature for 15 seconds (annealing of primer), and 72°C for 1 min and 30 seconds (extension of newly synthesised DNA strands). Reaction were concluded by heating at 72°C for 10 min and then held at 12°C. The T_m varied according to the length and composition of the primers used and the extension temperature varied according to the specific DNA polymerase used.

5.2.4.4 Visualisation of DNA

To visualise PCR and enzyme digest products, DNA was separated on an agarose gel at appropriate agarose content percentage in TAE buffer (40mM Tris Acetate and 1mM EDTA), containing GoldView Nucleic Acid Stain (Acridine orange; SBS Genetech Co., Ltd, Beijing, China) with appropriate DNA ladder and visualised under UV light.

5.2.4.5 PCR product purification

Prior to sequencing, PCR products were purified using Promega Wizard SV Gel and PCR Clean-Up System (Promega, https://www.promega.com.au/) with the microcentrifuge protocol.

5.2.4.6 Sequencing

Purified PCR fragments were sequenced by Macrogen (South Korea, https://dna.macrogen.com/eng/) or Australian Genome Research Facility (AGRF, http://www.agrf.org.au/). Sequencing results were aligned and analysed using Sequencher 4.0 (Genecodes, http://www.genecodes.com/).

5.2.5 Gene expression

5.2.5.1 RNA extraction

Harvested tissue was frozen in liquid nitrogen and homogenised by using tungsten carbide beads and mill mixer. Total RNA was extracted using RNeasy Plant Mini Kit (Qiagen, https://www.qiagen.com/us/shop/sample-technologies/rna/total-rna/rneasy-plant-mini-kit/#productdetails).

5.2.5.2 Quantification

RNA concentration was determined using NanoDrop 8000 Spectrophotometer (Thermo Scientific, https://www.thermofisher.com/order/catalog/product/ND-8000-GL)

5.2.5.3 Reverse transcription

First strand cDNA was synthesised from equal amount of (5µg) of total RNA using the SuperScript III Reverse Transcriptase (Thermo Fisher, https://www.thermofisher.com/), an engineered version of M-MLV RT with reduced RNase H activity and increased thermal stability. RNA was first incubated with 1µL of oligo(dT)$_{20}$ (50µM), 1µL 10mM dNTP mix and sterile water to a total of 13µL at 65°C for 5 min and then on ice for at least 1 min. A master mix was prepared (Table 5.4) and a 7µL aliquot was added to each tube. A negative control without reverse transcriptase (RT-) was included for each sample to check for the presence of genomic DNA contamination in the RNA sample. Reverse transcription was performed at 50°C for 30 - 60 min with a final incubation at 70°C for 15 min for enzyme inactivation.

Table 5.4 Reagents required for a 20µL reverse transcription reaction.

Reagent	**RT+**	**RT-**
RNA + oligo(dT)$_{20}$ (50µM) + 10mM dNTP mix + sterile water	13µL	13µL
First-Strand Buffer (5x)	4µL	4µL
DDT (0.1M)	1µL	1µL
RNaseOUT Recombinant RNase Inhibitor	1µL	1µL
Superscript III RT	1µL	(1µL of water)

5.2.5.4 Quantitative reverse transcription PCR (qRT-PCR)

For analysis of relative gene expression, qRT-PCR was conducted using a Rotor-Gene 3000 Real-time Thermal Cycler with Rotor-Gene 6 Version 6.1 (Corbett Research, Australia). A Corbett Robotics CAS-1200TM pipetting robot (Corbett Research, Australia) with CAS Robotics 4 Version 4.9.8 (1.6.61) software was used to prepare reactions. Each 10µL reaction comprised 2µL cDNA template, 5µL 2x Quantace SensiMixPlus SYBR reagent (Alexandria, NSW, Australia), 0.3µL each of forward and reverse primer (10µM) and 2.4µL autoclaved Milli-Q water. A no-template control (containing water instead of cDNA) was included for each run to check for contamination, and each sample was run in duplicate for increased accuracy. For each cDNA sample, actin was run on the reverse transcriptase negative

control (RT-) to check for contamination. Reactions were run for 50 cycles. A standard curve for the target gene was included in every run. Standard curves were generated from a 10-fold dilution series from 1 x 10-2 to 10-6 ng/µL. Standard curve regression was considered acceptable if the R2 value was equal to or higher than 0.99. Actin was chosen as the reference gene for evaluating transcript levels of flowering genes as previously described (Foo *et al*., 2005; Hecht *et al*., 2011; see Appendix 1 for primer details). Calculations of gene expression relative to actin were based on non-equal amplification efficiencies and deviation in threshold cycle using the means for two technical replicates.

5.2.6 TILLING

PsVHPO1 mutants were identified from an ethyl methanesulfonate mutant collection of 2,500 pea lines using TILLING screening as described by Dalmais *et al*. (2008). Details of all primer sequences are given in Appendix 1. The nature of the mutations was identified by sequencing. PsVHPO1 genomic sequence and TILLING mutations were integrated in UTILLdb (http://urgv.evry.inra.fr/UTILLdb).

5.2.7 Design and use of molecular markers for genotyping

5.2.7.1 Derived Cleaved Amplified Polymorphic Sequence (dCAPS) markers

dCAPS markers were designed using the dCAPS Finder 2.0 with a minimal number of mismatches (http://helix.wustl.edu/dcaps/dcaps.html) (Neff *et al.*, 2002) for SNPs which did not alter a restriction enzyme recognition site. Enzyme digests were conducted according to manufacturer's instructions (New England Biolabs, Inc., Ipswich, MA). Prospective dCAPS markers were tested on PCR products from parental lines and successful markers were used to genotype samples from the appropriate plant population(s) by standard PCR, restriction enzyme digest and visualisation of restriction enzyme products.

Psvhpo1 mutants were identified by the presence of 130-bp and hardly visible 30-bp fragments from primer pair PsHPO1-DraI-F and VHPO1-HRM-2R (see Appendix 1), which amplify a 160-bp PCR product from the PsVHPO1 gene, which was digested by DraI (New England Biolabs, Inc., Ipswich, MA) according to the manufacturer's

instructions, in homozygous wild-type plants. Heterozygous plants were identified by the presence of 160-bp, 130-bp and hardly visible 30-bp.

5.2.7.2 High Resolution Melt (HRM) markers

High Resolution Melt (HRM) analysis is a molecular technique for the detection of polymorphism in double-stranded DNA samples based on the temperature required for melting (separation) of two strands. HRM was used for genotyping of the plants and progenies of Family line 2760 and Family line 3023, using the Roto-Gene Q system (Qiagen), Roto-Gene Q System Software (Qiagen) and Roto-Gene ScreenClust HRM Software (Qiagen) (https://www.qiagen.com/au/).

5.2.8 Construction of alignments and phylogenetic trees

For phylogenetic analyses, amino acid sequences of proteins were aligned using ClustalX (Thompson *et al.*, 1997) and adjusted manually, where necessary, using GeneDoc (Version 2.7.000; http://www.psc.edu/biomed/genedoc) (Nicholas and Nicholas, 1997). Using these alignments, distance-based methods were used for phylogenetic analyses in PAUP* 4.0b10 (http://paup.csit.fsu.edu/). For comparison of homologous proteins, percentage identity at the amino acid level was calculated in GeneDoc from full-length protein alignments constructed using ClustalX.

5.2.9 Extract preparation for the detection and quantification of compounds

For the extraction and quantification of IAA and 4-Cl-IAA from young, fresh tissues, 0.3-2.5g of tissue was weighed (±0.0001 g FW), placed into a falcon tube with four volumes of cold (-20°C) extraction solvent (80% methanol in water (v/v), with BHT (250 mg/L). The tissue was then homogenised and held at 4°C overnight to extract. For quantification of these compounds, the supernatant added with [$^{13}C_6$] IAA (Cambridge Isotope Laboratories) and [D_2] 4-Cl-IAA (supplied by Prof. Jerry Cohen, Department of Horticultural Science, University of Minnesota). The samples were reduced under vacuum at 35°C, taken up in 2% acetic acid in water (v/v), and partitioned twice against diethyl ether. After drying the ether, samples were taken up in 1% acetic acid in water (v/v) and centrifuged for 5 min at 13 000 g. Aliquots were then taken for analysis by UPLC-MS/MS as described below previously (Tivendale

et al., 2012). The same aliquots were tested for methyl esters of IAA and 4-Cl-IAA by UPLC-MS/MS in this study.

For the extraction of Trp and 4-Cl-Trp, tissue was weighed, homogenised and extracted as described above, but distilled water with BHT (250mg/L) was the usual extraction solvent.

5.2.10 Statistical analysis

All statistical analyses were conducted using IBM® SPSS® Statistics Version 21, using a significance level of 0.05. Where sample size differed between groups, Levene's test for homogeneity of variance was used to test the assumption of equal variance. For comparisons between only two groups, two-tailed t-tests were conducted with equal variance either assumed or not assumed, depending on whether the assumption of equal variance had previously been rejected. For comparisons between three or more groups, one-way ANOVA was conducted. Where the assumption of equal variance was retained, a standard ANOVA was conducted to test for any significant differences between groups, and where significant differences were found, Tukey's HSD post-hoc test was conducted to further examine the differences. Where the assumption of equal variance was rejected, Welch's test was conducted to test for any significant differences between groups, coupled with the Games-Howell post hoc test if significant differences were found. p-values are reported in text for each t-test, standard ANOVA/Welch's test where no significance was found between groups, and post-hoc test (Tukey's HSD/Games-Howell) for specific comparisons where a standard ANOVA/Welch's test identified a significant difference between groups.

5.3 RESULTS AND DISCUSSION

A phylogram was constructed from a sequence alignment of inferred protein sequences for putative acid phosphatase/vanadium-dependent haloperoxidase-like genes of pea (Ps), *Medicago* (Mt), *Lotus* (Lj), *Cicer arietinum* (Ca), *Trifolium pratense* (Tp) and selected angiosperm species with conserved regions of high similarity (Figs. 5.2 and 5.3). Bootstrap values obtained from 1000 trees are indicated as a percentage above or next to each branch. Based on phylogenetic analysis among the species mentioned, PsVHPO1 formed a clade with Medicago VHPO-like gene (MtVHPO1) and *Trifolium pratense* (Tp1). PsVHPO2 and its clade formed a sister group to PsVHPO1 whereas PsVHPO3 and its clade are the most distant group from PsVHPO1. RT-PCR showed that the expression level of PsVHPO1 was approximately four times higher in whole seed (CT= 30.54, CT_{actin}=23.07) than apical shoot (CT=32.59, CT_{actin}=23.76) (Table 5.5). In the VHPO1 clade, all three species contained 4-Cl-IAA. Considering the phylogram, and the gene expression result by RT-PCR, PsVHPO1 was chosen to create a mutant population through TILLING. Five family lines of mutations were successfully created but only two lines (2760 and 3023) were useful with regards to their missense mutation that causes protein change, whereas the other three were less useful due to the silent or intronic nature of the mutation (Table 5.6). The *Psvhpo1* mutant family line 2760, with a missense mutation (G to A) causing protein change from glutamic acid to lysine, and family line 3203, also with a missense mutation (C to T) causing protein change from proline to serine, were chosen for further study. These lines were then backcrossed to the parental cv Caméor followed by subsequent selfings.

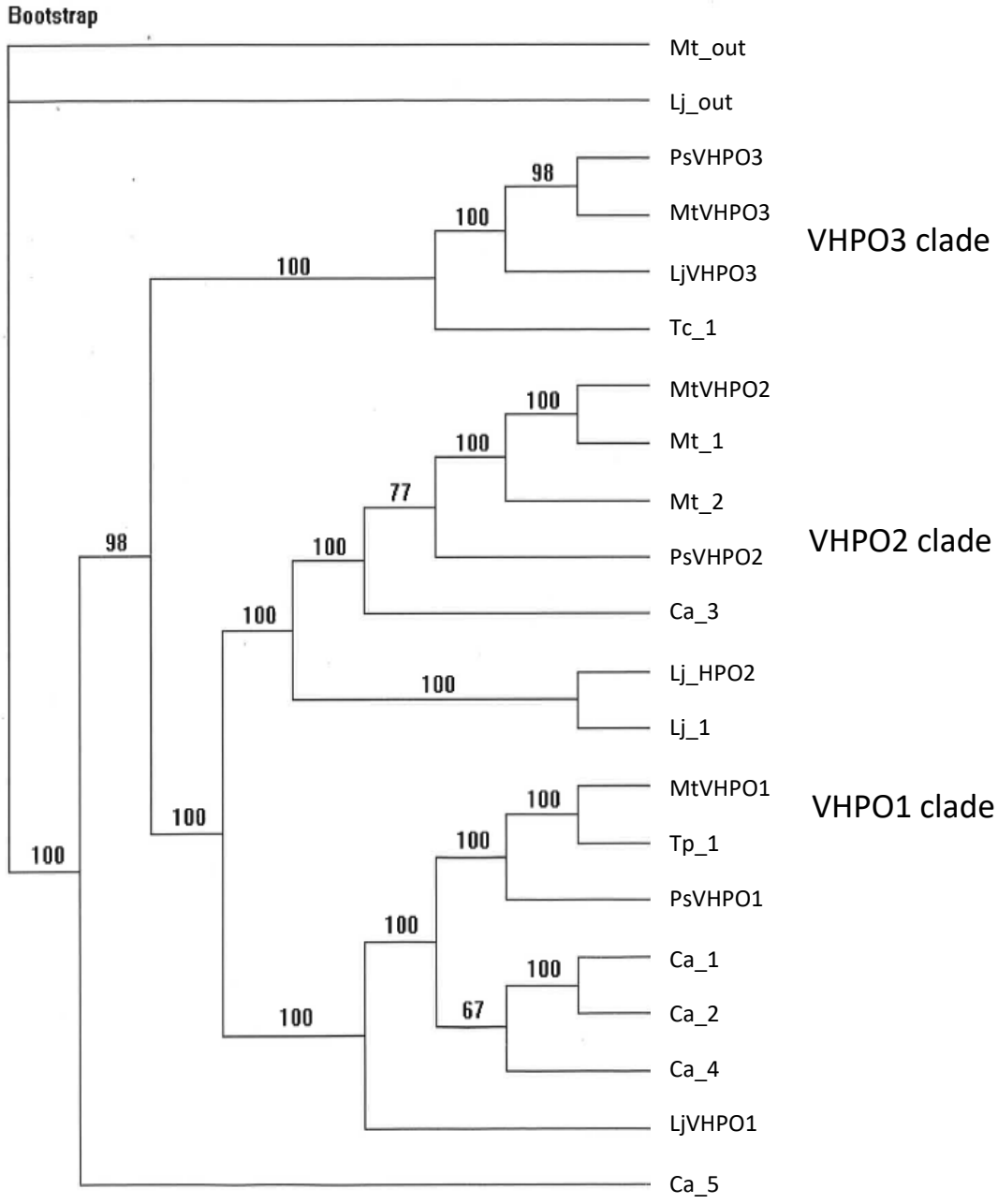


Figure 5.2 Phylogenetic neighbor-joining tree of the putative acid phosphatase/vanadium-dependent haloperoxidase-like protein sequences in legumes and selected angiosperm species (For abbreviation see Table 5.2).

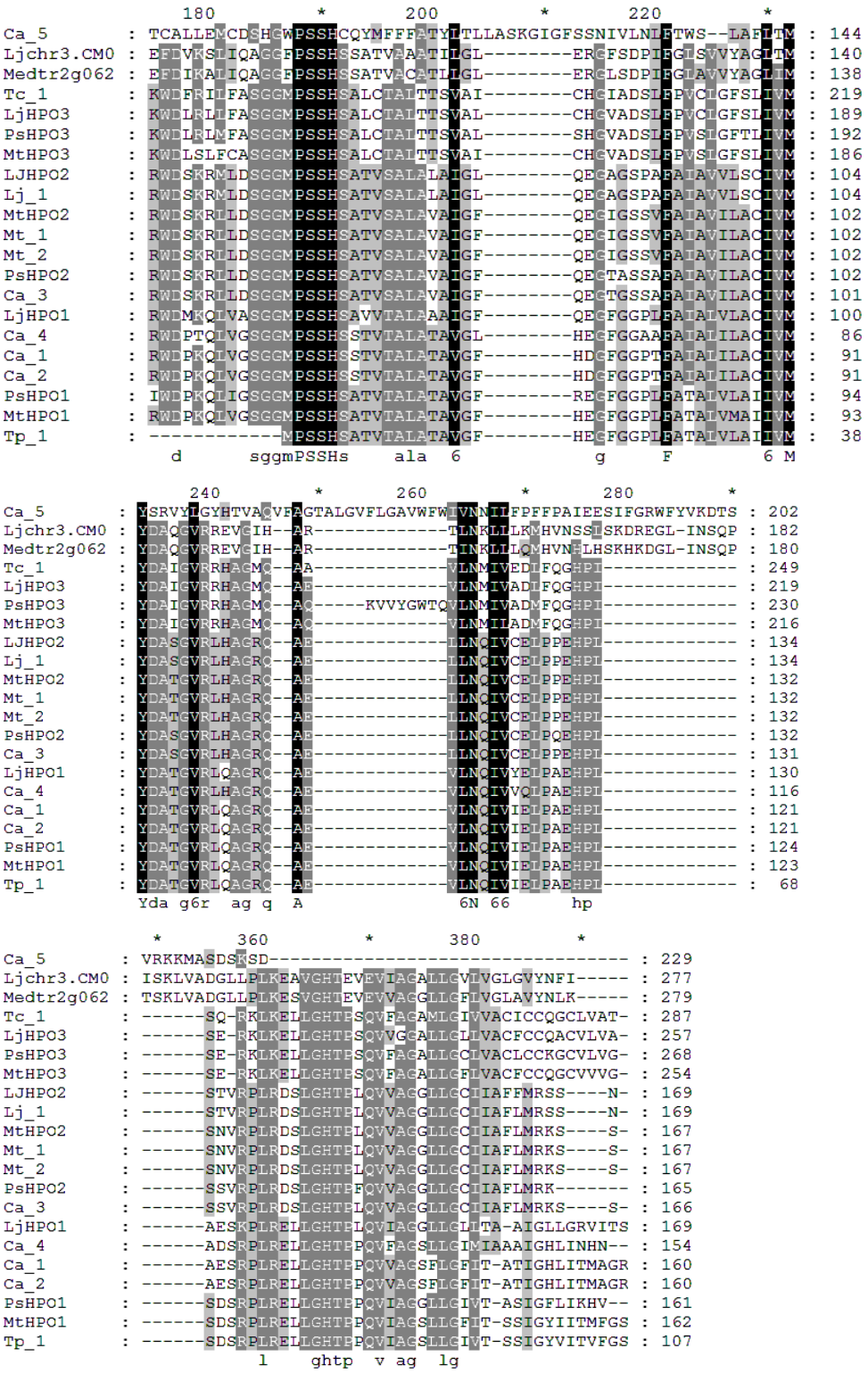


Figure 5.3 A sequence alignment of inferred protein sequences of putative VHPOs in pea, *Medicago*, *Lotus*, *Trifolium*, *Cicer* and *Cacao* aligned by Clustal Omega, showing conserved regions. Genedoc was used to shade the residues in the following manner: dark=100%, gray=75%, light gray=50% and white<50% identity. All accession number can be found in Table 5.2.

Table 5.5 CT values of the *PsVHPO1* gene and an actin-encoding gene in whole seeds and apical buds by Rt-PCR.

CT values		
Tissue type/Gene	**VHPO1**	**Actin**
Whole seed	**30.54**	**23.07**
Apical bud	**32.59**	**23.76**

Table 5.6 Details for mutant lines of the *PsVHPO1* gene obtained by TILLING. Point mutations are highlighted in colours for Line 2760 (green) and Line 3023 (yellow).

Ref	**Variant**	**Target Position**	**Type of mutation**	**gDNA position**	**cds position**	**Protein change**	**Line**	**zygosity**
C	T	228	MISSENSE	1725	412	P138S	3023	Htz
G	A	165	MISSENSE	1662	349	E117K	2760	Htz
G	A	221	SILENT	1718	405	G135G	3077	Htz
G	A	67	INTRONIC	1564	-	-	4109	Htz
G	A	339	INTRONIC	1836	-	-	4263	Htz

Coding sequence of *PsVHPO1*

ATGGACGACGCTGCAGGTACAACCACGTCTTCTTCTTCTTCAACCTTTGGCAATTACCCTCTTTTTTGCGCCATA
GTTGCTTTTACCATCGCTCAACTCATCAAATTCTTCACCGCTTGGTATAAGGAAAGGATATGGGATCCGAAACA
ATTGATTGGATCTGGCGGAATGCCGTCTTCTCATTCAGCTACTGTTACTGCTCTTGCTACAGCGGTTGGATTTCG
GGAAGGATTTGGAGGACCACTTTTCGCTACTGCATTGGTTCTGGCTATCATTGTGATGTATGATGCTACTGGTG
TAAGATTGCAAGCAGGACGACAAGCGGAGGTTCTTAATCAAATTGTAATTGAACTTCCTGCTGAACATCCTCTG
TCTGACAGCAGACCTCTTCGCGAACTTCTTGGGCATACCCCGCCTCAGGTAATTGCTGGTGGTTTACTTGGAAT
CGTAACAGCATCTATTGGTTTTTTAATAAAACATGTTTGA

Protein sequence of *PsVHPO1*

MDDAAGTTTSSSSSTFGNYPLFCAIVAFTIAQLIKFFTAWYKERIWDPKQLIGSGGMPSSHSATVTALATAVGFREGF
GGPLFATALVLAIIVMYDATGVRLQAGRQAEVLNQIVIELPAEHPLSDSRPLRELLGHTPPQVIAGGLLGIVTASIGFL
IKHV

The heterozygous, segregating nature of both lines 2760 and 3023 was genotyped by HRM and/or dCAPs markers, and then confirmed by sequencing. The flowers of Caméor mother pea plants were hand-pollinated with pollen of heterozygous (HET) plants. As documented in Figure 5.4, in the first generation of the cross, a 1:1 WT:HET ratio was expected. The heterozygous progenies were sown as the second generation to be selfed, expecting a 1:2:1 WT:HET:mutant ratio but none of the progenies of known heterozygous plants contained mutants. However, another known heterozygous plant did produce 2 mutants (x2760_02_12_08_06 and x2760_02_12_08_11), out of 40 plants, as genotyped by dCAPs marker and digest (Fig. 5.5). Out of sixty seeds produced from another count of selfing of heterozygous plants, one mutant (x2760_2_13_09) was generated. WT and heterozygous clusters were shown by PCA and HRM but the mutant was grouped to the heterozygous cluster (Fig. 5.6). The mutant was only then revealed by dCAPS markers (Fig. 5.7) and sequencing (Fig. 5.8).

Such a low proportion of mutants is unusual and therefore Line 3023 (heterozygous) was tested. Similarly, 24 seeds from Line 3023_04 again did not conform to 1:2:1 WT:HET:mutant ratio as there was only one mutant generated, while there were 10 WT and 13 HETs, as revealed by HRM (Fig. 5.9) then confirmed by sequencing (Fig. 5.10). HRM was very effective for separating HETs from WT seedlings but not for separating HETS from mutants, in which the latter was done by sequencing. On the other hand, dCAPs markers are equally effective as HRM for screening large populations so that, without getting all samples sequenced, sequencing can target samples of interest for confirmative genotyping.

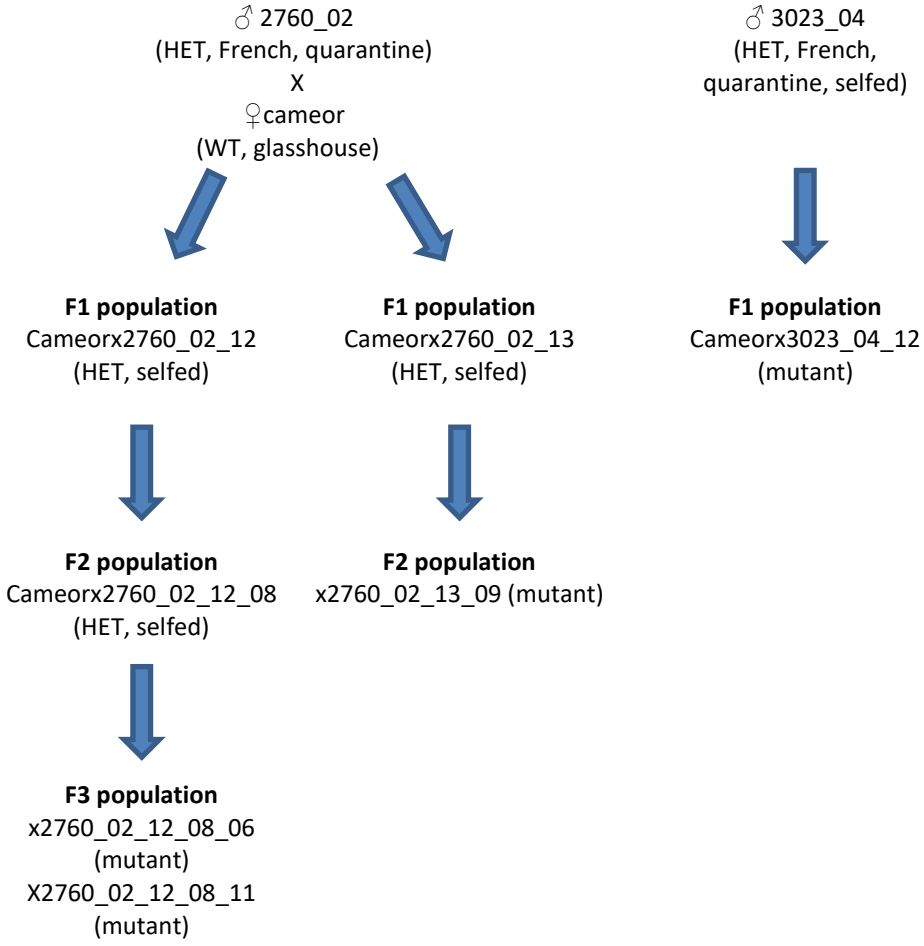


Figure 5.4 Family tree of mutation families 2760 and 3023 of *PsVHPO1*. Cameorx2760_02_12 is F_1 plant (i.e. heterozygous) but did not produce any mutant seeds. The selfing of cameorx2760_02_12_08, F_2 plant produced two mutant seeds (x2760_02_12_08_06 and x2760_02_12_08_11). Another F_1 plant cameorx2760_02_13 was subjected to selfing and successfully produced one mutant seed (x2760_02_13_09). The selfing of cameorx3023_04 produced one mutant seed (x3023_04_12).

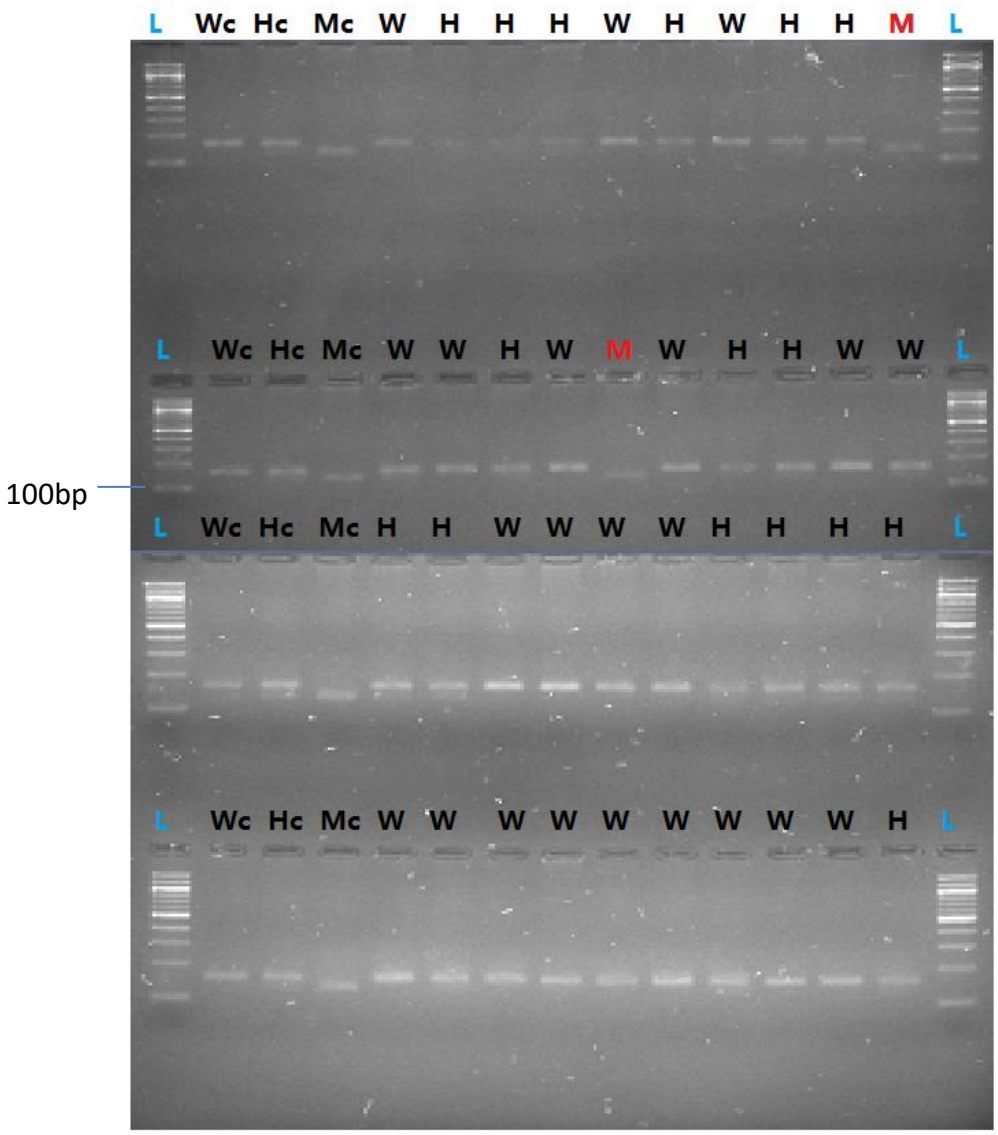


Figure 5.5 Digest for *PsVHPO1* gene of 40 seedlings whereas restriction fragments of x2760_02_12_08_06 and x2760_02_12_08_11 were visually similar to mutant control. L represents 100bp base pair NEB ladder; Wc, wildtype control; Hc, heterozygous control; Mc, mutant control; W, wildtype (WT); H, heterozygous; M, mutant.

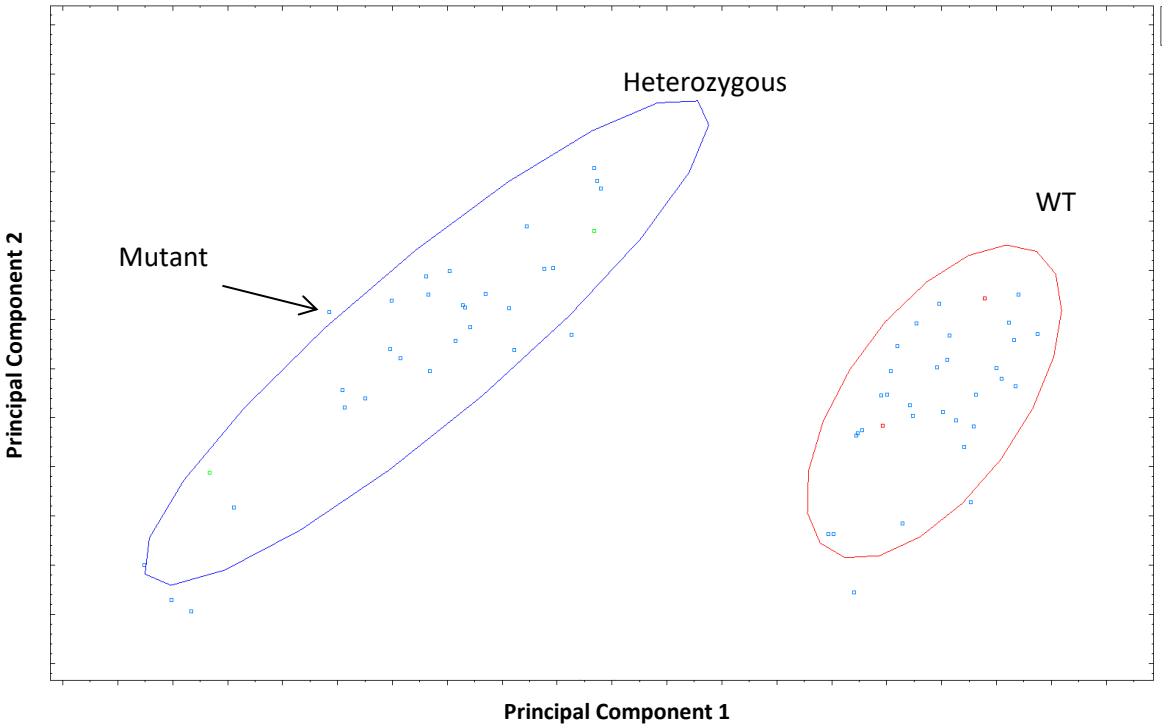


Figure 5.6 Genotyping by HRM and Principle Component Analysis showing clusterings of WT (red circle, red square indicates WT control), heterozygous (blue circle, green square indicates heterozygous control) and mutant x2760_02_13_09 (noted by arrow) seedlings from 60 seeds produced by selfing of heterozygous pea plants. Mutant and heterozygous seedlings were confirmed by sequencing.

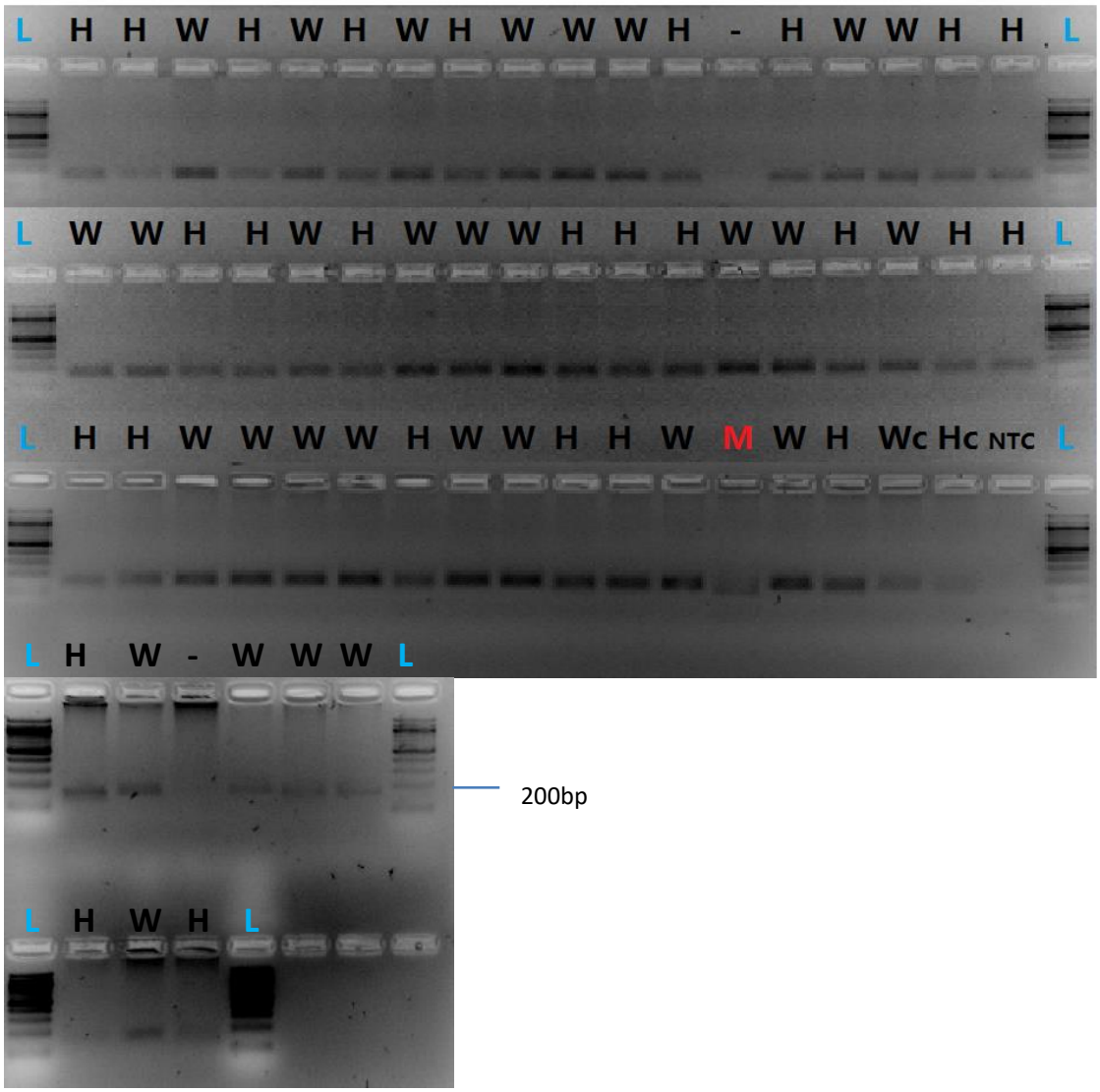


Figure 5.7 Digest for *PsVHPO1* gene of 60 seedlings whereas restriction fragments of x2760_02_13_09 is visually different to the rest. L represents 100bp base pair NEB ladder; W, wildtype (WT); H, heterozygous; M, mutant; Wc, wildtype control; Hc, heterozygous control; NTC, no template control.

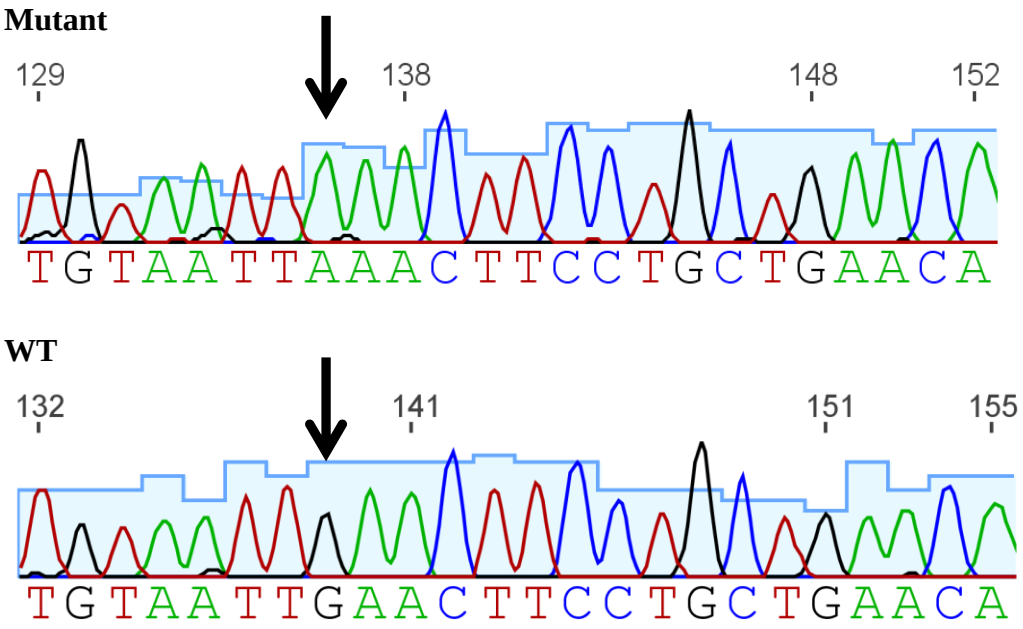


Figure 5.8 A single point mutation of mutant seedling x2760_2_13_09 was confirmed by sequencing as compared to WT seedling . Black arrow points out the G to A nucleotide change.

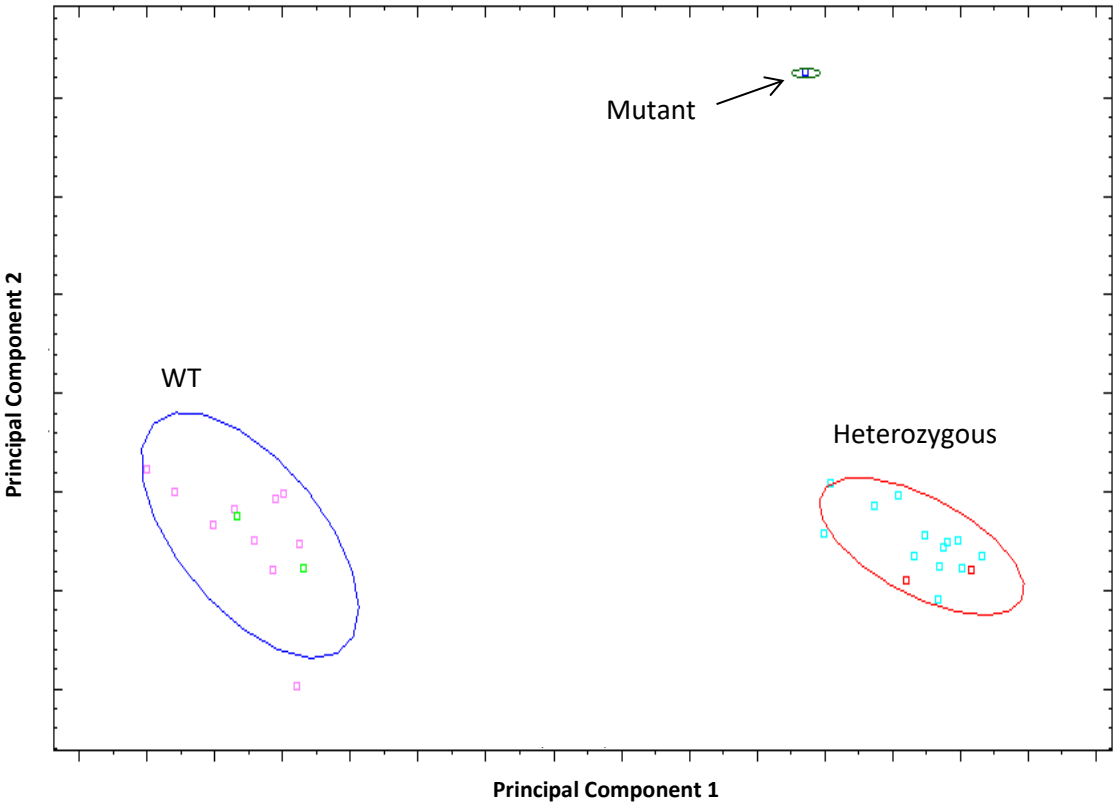
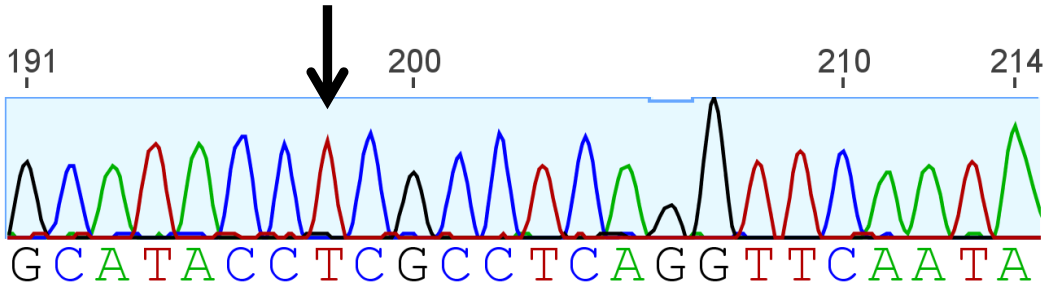


Figure 5.9 Genotyping by HRM and Principle Component Analysis showing clusterings of WT (blue circle, green square indicates WT control), heterozygous (red circle, red square indicates heterozygous control) and mutant x3023_04_12 (blue square, noted by arrow) seedlings from 60 seeds produced by selfing of heterozygous pea plants. Mutant and heterozygous seedlings were confirmed by sequencing.

Mutant



WT

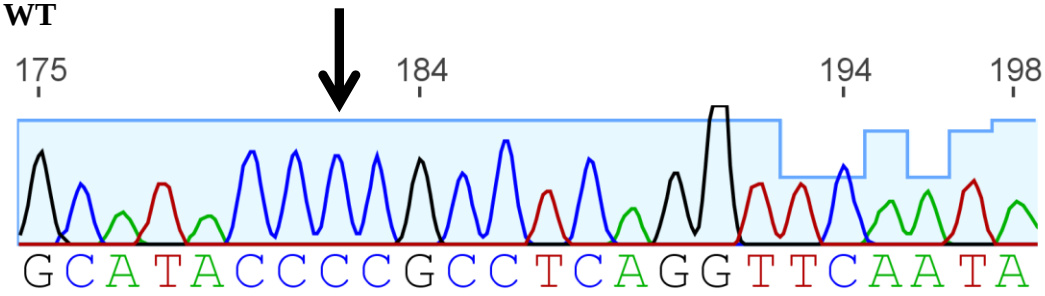


Figure 5.10 A single point mutation of mutant seedling x3023_04_12 was confirmed by sequencing as compared to WT seedling. Black arrow points out the C to T nucleotide change.

The visual seed phenotype of all three mutants was normal (Fig. 5.11). This prompted the quantification of hormonal levels in both mutant and WT seeds. However, no reduction of hormonal levels of both IAA and 4-Cl-IAA was observed in *PsVHPO1* mutants compared to the WT. In fact, there was more 4-Cl-IAA in young mutant seeds than young WT seeds but the reason for that observation is not known. Interestingly, although internal standards for methyl esters of IAA (IAAme) and 4-Cl-IAA (4-Cl-IAAme) were not added, the signal ratios of IAAme:IAA and 4-Cl-IAAme:4-Cl-IAA were reduced in young mutant seeds (Fig. 5.12). The presence of methyl esters of IAA and 4-Cl-IAA in the young mutant seeds are real as the detection is based on previously known retention times and fragmentation pattern of the standards of those compounds. The levels of these compounds, however, cannot be determined as there was no labelled internal standards added in the samples. This may indicate that mutation in *PsVHPO1* may have caused a slight reduction of 4-Cl-IAAme at the very early developmental stage as observed in the young seeds.

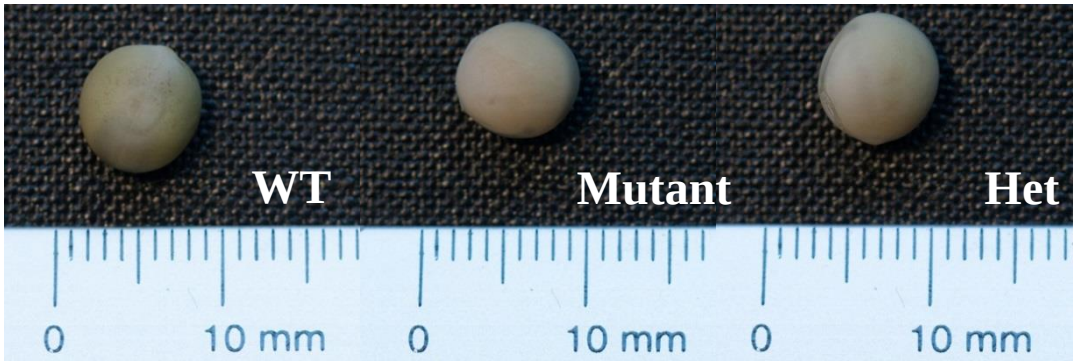
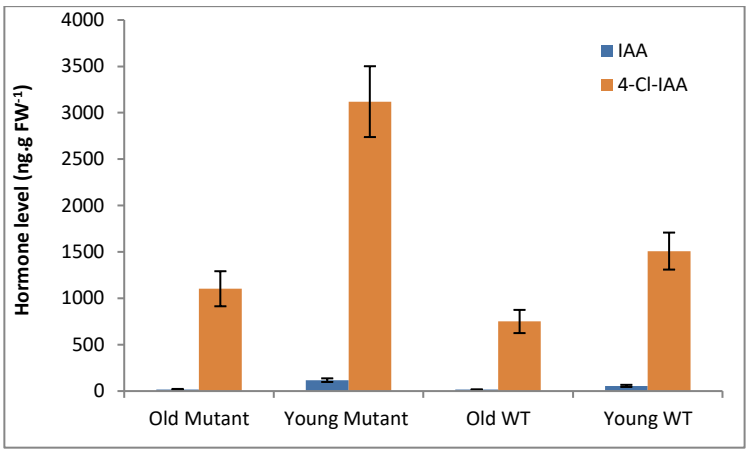


Figure 5.11 Phenotypes of wildtype (WT), mutant and heterozygous (Het) seeds of Line 2760.

A



B

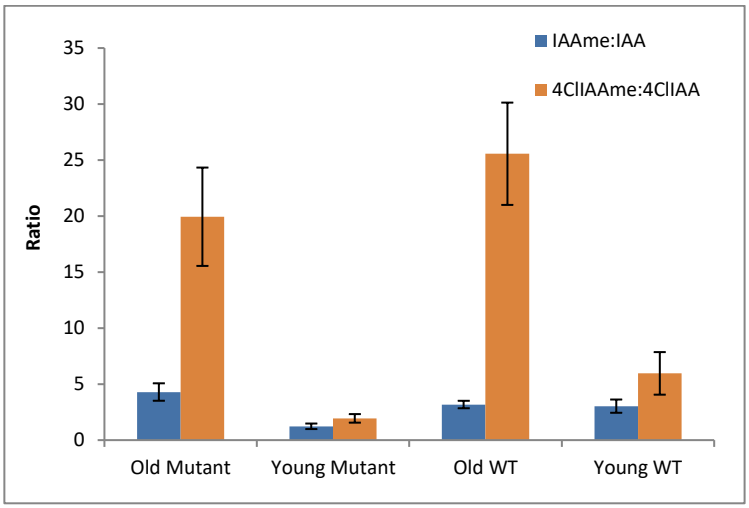


Figure 5.12 (A) IAA and 4-Cl-IAA levels and (B) IAAme:IAA and 4-Cl-IAAme:4-Cl-IAA signal ratios in the seeds of WT and mutant at young and old stages. Bars are means ± standard errors (n = 6 for IAA and 4-Cl-IAA as well as grouping ratios of IAAme:IAA and 4-Cl-IAAme:4-Cl-IAA).

5.3.1 Haloperoxidase or not?

One explanation for no reduction in 4-Cl-IAA in the PsVHPO1 mutant seeds is that the phenotypic effect may be rescued by redundant genes encoding other members of the pea VHPO family. However, another explanation is that the chosen gene may not be responsible for halogenation at all. The chosen gene, VHPO1 as well as the other two unstudied VHPO2 and VHPO3, were annotated as acid phosphatase/vanadium-dependent haloperoxidase-like genes in the online Pea RNA-Seq gene atlas. While the known VHPOs from brown algae, red algae, fungi and bacteria (Table 5.7) do not share a high level of overall protein sequence identity, they do contain two short, highly conserved domains related to vanadate binding, which are (1) PxYxSGHA and (2) LxxxxAxxRxxxGxHxxxD (Leblanc *et al.*, 2015) (Fig. 5.13). A BLAST search using those known bacterial and fungal VHPO protein sequences did not find any significant alignment with sequences from pea, *Medicago truncatula* and *Lotus japonicas* (Fig. 5.14). Therefore, the annotation may be informatively limited, or even incorrectly assigned. It is possible, therefore, that the studied PsVHPO1 (and PsVHPO2 and PsVHPO3 as well) do not possess a halogenating function. This prompted a revisiting of other halogenating enzymes as candidate(s) for plant halogenases.

Table 5.7 Vanadium-dependent haloperoxidases

Organism and species		Halogenating activity	Genbank accession #	Abbreviated code
Brown algae	*Ascophyllum nodusum*	VBPO	P81701	An1_VBPO
		VBPO	CCD42013	An2_VBPO
	Laminaria digitata	VBPO	CAD37191	Ld_VBPO
		VIPO	CAF04025	Ld_VIPO
Red algae	*Corallina officinalis*	VBPO	AAM46061	Co_VBPO
	Corallina pilulifera	VBPO	BAA31261	Cp1_VBPO
			BAA31262	Cp2_VBPO
	Gracilaria changii	VBPO	AGE00855	Gc_VBPO
Fungi	*Curvularia inaequalis*	VCPO	CAA59686	Ci_VCPO
	Alternaria didymospora	VCPO	CAA72622	Ad_VCPO
	Zobellia galactanivorans	VIPO	YP_004735706	Zg_VIPO
Bacteria	*Rhodopirellula baltica SH1*	VCPO	CAD72609	Rb_VCPO
	Streptomyces sp. CNQ525	VCPO	ABS50486	Ss1_VCPO
	Streptomyces sp. CNH189	VCPO	AGH68925	Ss2_VCPO

		AGH68909	Ss3_VCPO
	Synechococcus sp. WH 8020	VBPO	ADO20318	Ss4_VBPO

VBPO: vanadium-dependent chloroperoxidase; VCPO: vanadium-dependent chloroperoxidase; VIPO: vanadium-dependent iodoperoxidase

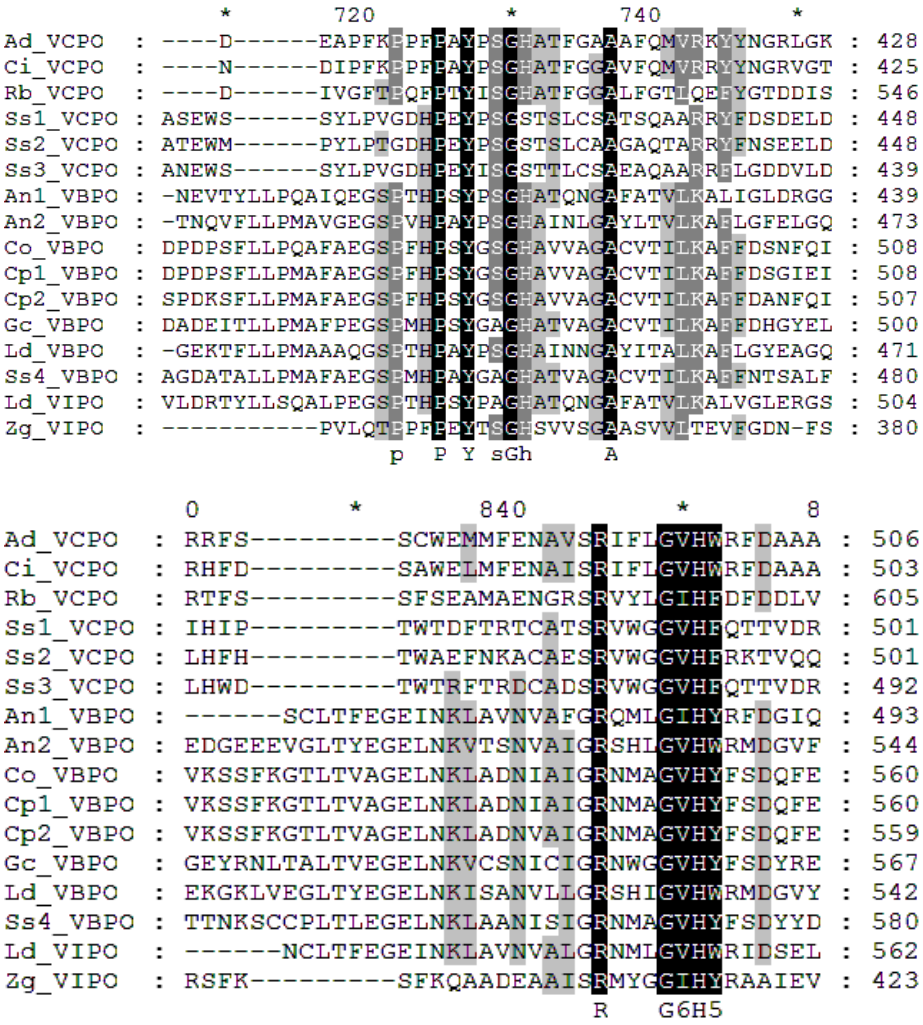


Figure 5.13 Multiple amino acid sequence alignment of VHPOs (VCPOs, VBPOs and VIPOs) identified in brown algae, red algae, fungi and bacteria aligned by Clustal Omega. Gendoc was used to shade the residues in the following manner: dark=100%, gray=75%, light gray=50% and white<50% identity. All abbreviated codes and accession numbers for the enzymes can be found in Table 5.7.

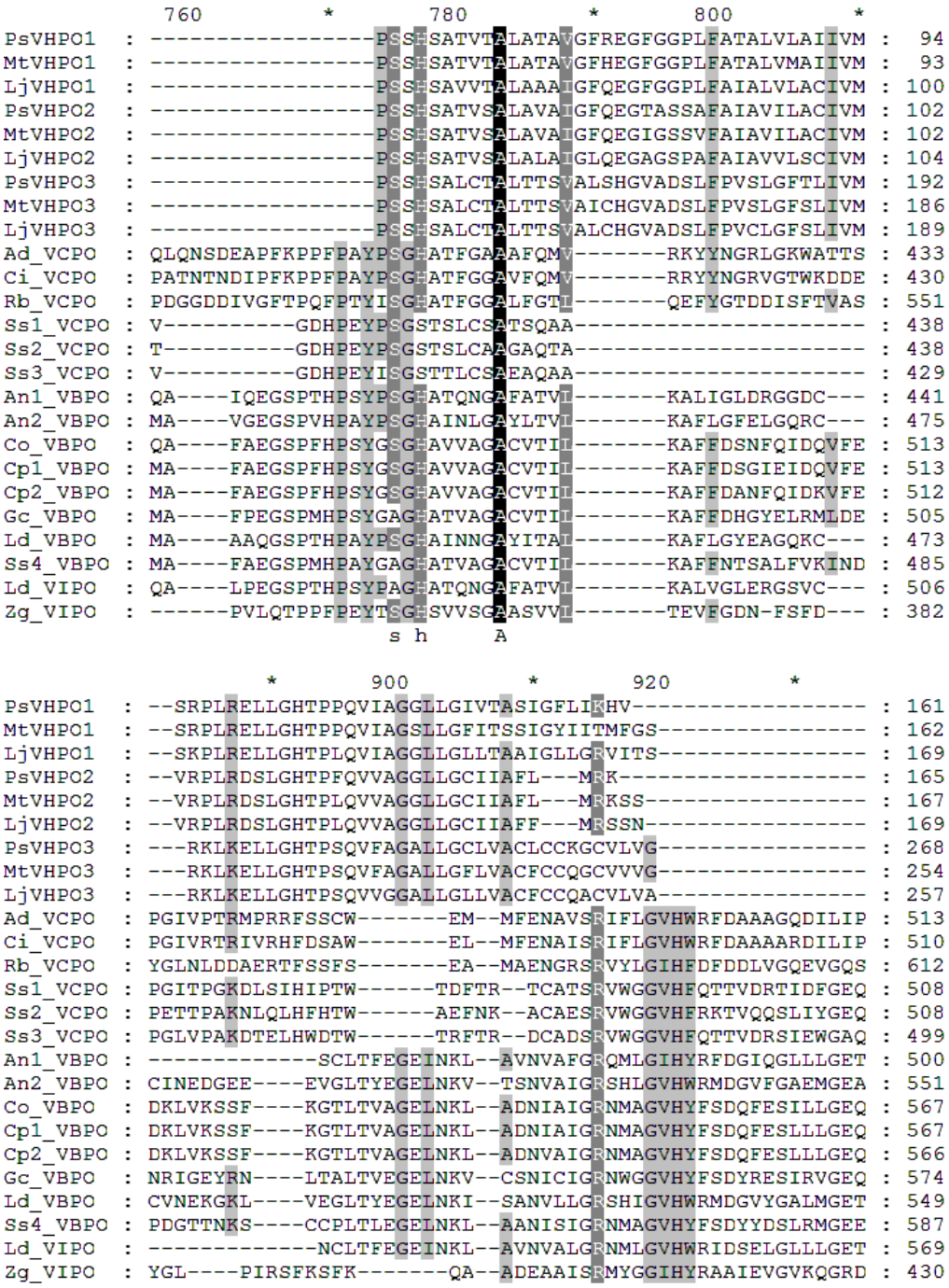
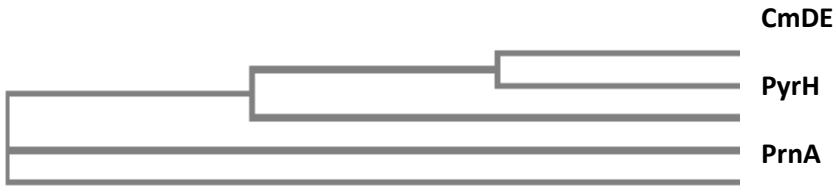


Figure 5.14 Multiple amino acid sequence alignment between putative VHPOs in legumes and VHPOs (VCPOs, VBPOs and VIPOs) identified in brown algae, red algae, fungi and bacteria aligned by Clustal Omega. Gendoc was used to shade the residues in the following manner: dark=100%, gray=75%, light gray=50% and white<50% identity. All abbreviated codes and accession numbers for leguminous

VHPOs were documented in Table 5.2 whereas VCPOs, VBPOs and VIPOs for non-leguminous species were recorded in Table 5.7.

Investigation of another group of halogenating enzymes, the flavin-dependent halogenases revealed that five of these enzymes that use tryptophan have a few conserved regions (Fig. 5.15) but do not have significant alignment with sequences from *Medicago*, pea and *Arabidopsis* in a BLAST search.

(A)



(B)

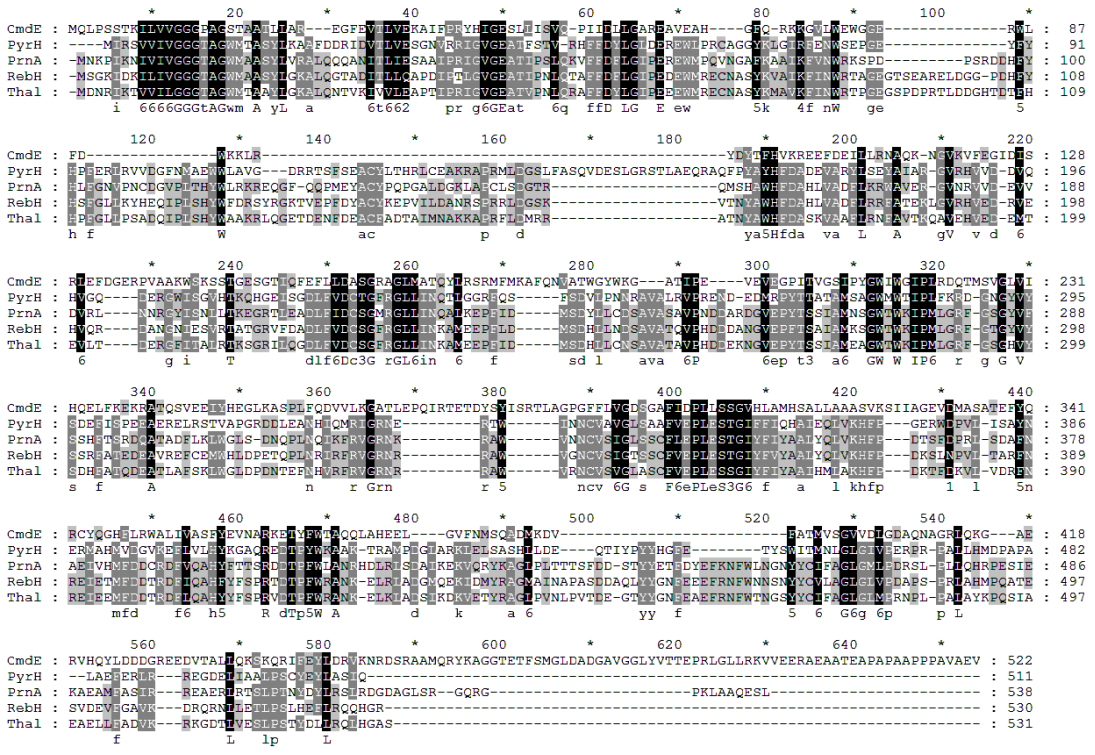


Figure 5.15 (A) Phylogenetic relationship of characterised flavin-dependent halogenases in bacteria and (B) amino acid sequence alignment of flavin-dependent halogenases in bacteria aligned by Clustal Omega. Gendoc was used to shade the

residues in the following manner: dark=100%, gray=75%, light gray=50% and white<50% identity. All abbreviated codes and accession numbers can be found in Table 5.1.

Another class of bacterial halogenating enzymes are the non-heme, iron-dependent halogenases, for example CytC3 and SyrB2 (Wong *et al.*, 2009). However, hydroxylation is the most common reaction catalysed by this enzyme superfamily. Interestingly, iron coordination by a 2-His-1-carboxylate motif with aspartate (D) (i.e. H-X-D) is typical for hydroxylases whereas halogenases have an alanine (A) substitution instead (i.e. H-X-A) (Wong *et al.*, 2009). In fact, there are hydroxylases, carboxylate oxidases, oxygenases or similar enzymes in plants that share high identity and homology with non-heme iron dependent halogenases. Therefore, plant halogenases may have evolved from hydroxylases, with a single mutation causing amino acid substitution D to A in the active site. Neumann *et al.* (2008) were "tempted" to postulate that the loss of iron-coordinating carboxylate ligand and coordination of exogenous halide may have given rise to halogenation from hydroxylation activity. It should be realised, however, that the attempts to convert halogenase to hydroxylase and vice versa have thus far failed by substituting the alanine, aspartate and glutamate in the 2-His-1-carboxylate motif.

A search in the literature has revealed another group of potential candidates, the cofactor-free HPO family enzymes (Xu and Wang, 2016). To date, five crystal structures of these bacterial cofactor-free HPO have been well studied, showing a S-H-D catalytic triad and great similarity to α/β hydrolase (Hecht *et al.*, 1994; Hofmann *et al.*, 1998). In fact, α/β hydrolases and orthologs are relatively abundant in *Arabidopsis*, *Medicago*, rice and petunia. Therefore, the pea halogenases may have evolved from those similar enzymes with a α/β hydrolase fold.

The discovery of non-chlorinating *Trigonella* in the previous chapter may indicate a readily exploitable natural system to study plant halogenation. For instance, by crossing closely related chlorinating and non-chlorinating *Trigonella* species the number of genes involved in the halogenation of Trp could possibly be revealed. If it is a 3 WT: 1 mutant ratio, it may indicate a single gene is responsible for this halogenation step in the legumes.

Chapter 6 General Discussion and Conclusion

The number of phytohormones known to science is not large and therefore the basic understanding of any plant hormone is very important at this moment in human history, as we face a possible food crisis due to increasing global population coupled with the consequences of global warming.

The nature of chlorinated auxin and its chlorinated precursors, which are unique to legumes and, in particular, their reproductive tissues (the seeds and seed pods, often the main part of food crops for human consumption) highlights that elucidating halogenation in plants is of paramount importance. Such key knowledge can potentially unlock the full capacity of these naturally nodulating food crop species to "feed the world" as well as play a vital part in nutrient cycles.

This thesis documents the extended understanding of chlorinated auxin and other precursors, such as tryptophan, through investigation of the distribution and evolutionary aspect of this halogenated hormone in planta, as well as a molecular study of a putative pea gene with a possible role in halogenation, a pea vanadium-dependent haloperoxidase. In the present study, the distribution of chlorinated auxin was revisited, by screening a large range of species spanning many genera in the Fabaceae for 4-Cl-IAA and 4-Cl-Trp using UPLC-MS (Chapter 2). A screening approach analysing half of the samples with labelled standards but the other half without the labelled standards has successfully confirmed the presence or absence of the hormones of interest by nullifying any ambiguity resulting from unlabelled 4-Cl-IAA (6%) in the labelled internal standard. The result overturned the old notion that 4-Cl-IAA is restricted to the tribe *Fabeae*, which used to be called *Viceae*, that includes *Vicia*, *Pisum*, *Lathyrus*, *Lens* and *Vavilovia*. In fact, 4-Cl-IAA and 4-Cl-Trp are present in another five leguminous genera, including *Trifolium*, *Melilotus*, *Trigonella*, *Medicago* and *Ononis*. The previously-reported presence of 4-Cl-IAA in the Scots pine was challenged in this study; there were no detectable levels of 4-Cl-IAA in that species, nor in four other *Pinus* species. This supports that 4-Cl-IAA and its chlorinated precursor(s) are unique to the Fabaceae in the entire kingdom Plantae. Nevertheless, 4-Cl-IAA and 4-Cl-Trp have been shown to be mainly restricted to the reproductive tissues (i.e. the seeds) rather than the vegetative tissues (as previously

reported in *Vicia faba* by Pless *et al.* (1984)), when leaf and seed samples of *Vicia faba* and *Trifolium repens* were analysed in this study.

The discovery of chlorinated auxin in *Trifolium*, *Melilotus*, *Trigonella*, *Medicago* and *Ononis*, for the first time is extremely significant because many species of these five genera are important agricultural crops around the world, grown as food crops or pasture species as well as commercial crops with high potential and values in nutraceautical industry. This study reported a total of 74 legume species with chlorinating ability, not known previously to that capacity (Fig. 6.1).

Apart from the chlorinating species, a range of species from other genera in the Fabaceae were analysed and shown to be non-chlorinating, including *Cicer*, *Galega*, *Parochetus*, *Astragalus* and *others*. These genera belong to other tribes or clades in the Fabaceae and were chosen based on their phylogenetic relationship to the chlorinating genera and tribes. For the first time, this study elucidates the evolution of the chlorinating ability, which appeared to have started approximately 25 million years ago, after the divergence of the non-chlorinating genera *Cicer* and *Galega* (Chapter 3). Plant phylogeny is never straightforward, and it will still take molecular geneticists and taxonomists a long time to resolve their agreements and disagreements. In this context, the unique chlorinating ability described here may be phylogenetically informative.

As previously reported, IAA is not converted into 4-Cl-IAA but, instead, Trp was chlorinated and converted into 4-Cl-IAA subsequently (Tivendale *et al.*, 2012). The screening in this study revealed that there may possibly be two losses of chlorinating ability in the evolution of *Trigonella* species (Chapter 4). The reason(s) behind this is still unknown but one group of taxonomically closely related *Trigonella* species, subgenus *Foenum-graecum*, as classified by both traditional and molecular classifications, appears to include chlorinating and non-chlorinating species. As documented in Chapter 4, *Trigonella foenum-graecum* (or fenugreeks) and *T. gladiata* do not produce either 4-Cl-IAA or 4-Cl-Trp, whereas the other species of the subgenus, *T. macrorrhyncha*, does chlorinate. This loss of chlorinating ability was only found in *Trigonella*. In the future, the chlorinating *T. macrorhynncha* could be crossed with either or both of the two non-chlorinating species to shed light

on how many chlorination genes are involved, based on the segregation ratio of chlorinating and non-chlorinating progenies of that cross. This approach, using an existing natural system (*Trigonella*) might provide a considerably more direct route to investigate halogenation in plants, as the closely related chlorinating and non-chlorinating species can be studied in parallel using physio-chemical methods and molecular tools. As far as interspecific fertility in *Trigonella* is concerned, very limited but some hybridisation works have been done. Singh (1973) reported successful crosses of *T. corniculata* with *T. hamosa* and *T. cretica* as well as crossing between the latter two species. However, fenugreek (*T. foenum-graecum*) and *T. coerulea* could not be crossed with any species mentioned earlier (Singh, 1973).

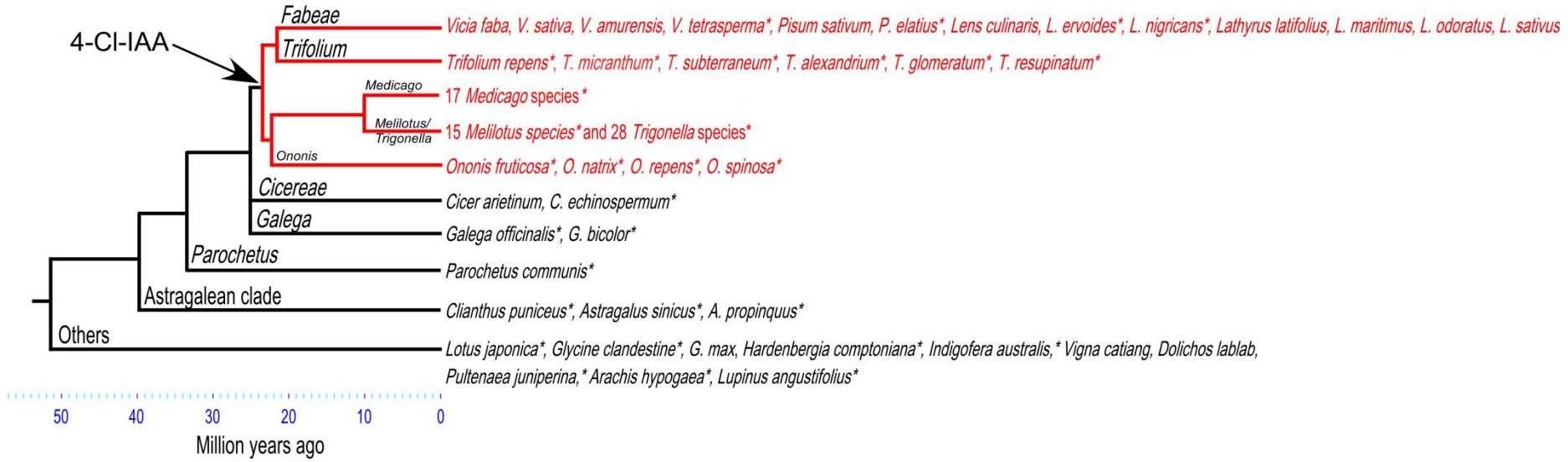


Figure 6.1 Phylogeny of the major lineages of the inverted-repeat-loss clade (IRLC) of the family Fabaceae with a particular focus on the closest lineages to the tribe *Fabeae*. The ability to produce 4-Cl-IAA appears to have evolved after the divergence of the genera *Cicer* and *Galega* approximately 25 million years ago (indicated by arrow and red branches). Species with detectable 4-Cl-IAA are shown in red; species that do not have detectable 4-Cl-IAA are shown in black. Species investigated in this study are indicated by an asterisk. Phylogenetic relationships and divergence dates are taken from Choi *et al.* (2004), Lavin *et al.* (2005), Schaefer *et al.* (2012) and Wojciechowski *et al.* (2004).

The investigation of a putative pea vanadium-dependent haloperoxidase showed that the studied mutants of PsVHPO1 generated through TILLING not only appeared normal visually but also produced chlorinated auxin throughout seed development (Chapter 5). PsVHPOs and putative VHPOs from *Medicago* and *Lotus* did not have good alignment to flavin-dependent halogenases that are known to chlorinate tryptophan at various positions in the indole ring in lower organisms except at position C-4. This finding prompted the reinvestigation of other types of halogenases in the form of bioinformatics by aligning sequences of known halogenating active sites with genetic sequences from plants, particularly *Medicago truncatula* or pea as well as *Arabidopsis thaliana*. For instance, non-haem iron dependent halogenases did not show good sequence alignment with *Medicago truncatula*. Blast result, however, revealed good similarity between those of *Medicago truncatula* (mainly α/β hydrolases) and cofactor free haloperoxidases of bacterial origin. This shows that a pea homolog of α/β hydrolases can be good candidate for future study as mutation(s) in these enzymes may have phenotypic effects in the chlorination process or ability in pea or any chlorinating legumes. In fact, the current availability and accessibility of sequenced genomes of closely related, non-chlorinating (*Cicer*, *Glycine*, *Lotus*) or chlorinating (*Pisum, Trifolium* and *Medicago*) leguminous species sharing reasonable level of synteny allow comparison among these species targeting both hydroxylases and α/β hydrolases genes.

In conclusion, this study improves the knowledge of the distribution and evolutionary origin of chlorination in plants. The discovery of chlorinating and non-chlorinating species in the genus *Trigonella* may provide a very useful natural system to study the genetic control of chlorination in plants. Future characterisation of α/β hydrolase gene homologs will also potentially contribute to the understanding of this process complementing the *Trigonella* approach. Addressing these research goals will be fundamental for improving our understanding of seed development, and may lead to improved seed crop breeding, especially in legumes, for better food reproduction and food security.

References

Agarwal V, Miles ZD, Winter JM, Eustáquio AS, El Gamal AA, Moore BS. 2017 Enzymatic halogenation and dehalogenation reactions: pervasive and mechanistically diverse. Chemical Reviews **117**, 5619-5674.

Alves‐Carvalho S, Aubert G, Carrère S, Cruaud C, Brochot AL, Jacquin F, Klein A, Martin C, Boucherot K, Kreplak J. 2015 Full‐length de novo assembly of RNA‐seq data in pea (Pisum sativum L.) provides a gene expression atlas and gives insights into root nodulation in this species. The Plant Journal **84**, 1-19.

Basu MM, González-Carranza ZH, Azam-Ali S, Tang S, Shahid AA, Roberts JA. 2013 The manipulation of auxin in the abscission zone cells of Arabidopsis flowers reveals that indoleacetic acid signaling is a prerequisite for organ shedding. Plant Physiology **162**, 96-106.

Butler A, Sandy M. 2009 Mechanistic considerations of halogenating enzymes. Nature **460**, 848-854.

Cernay C, Pelzer E, Makowski D. 2016 A global experimental dataset for assessing grain legume production. Scientific Data **3**, 160084.

Chaudhary L, Sanjappa M. 1998 Parochetinae: a new subtribe of Trifolieae (Leguminosae, Papilionoideae). Taxon **47**, 829-831.

Dalmais M, Schmidt J, Le Signor C, Moussy F, Burstin J, Savois V, Aubert G, Brunaud V, de Oliveira Y, Guichard C. 2008 UTILLdb, a Pisum sativum in silico forward and reverse genetics tool. Genome biology **9**, R43.

Dangi R, Tamhankar S, Choudhary RK, Rao S. 2016 Molecular phylogenetics and systematics of Trigonella L. (Fabaceae) based on nuclear ribosomal ITS and chloroplast trnL intron sequences. Genetic Resources and Crop Evolution **63**, 79-96.

De Jong M, Mariani C, Vriezen WH. 2009 The role of auxin and gibberellin in tomato fruit set. Journal of Experimental Botany **60**, 1523-1532.

Dharmasiri N, Dharmasiri S, Estelle M. 2005 The F-box protein TIR1 is an auxin receptor. Nature **435**, 441-445.

Engvild KC. 1975 Natural chlorinated auxins labelled with radioactive chloride in immature seeds. Physiologia Plantarum **34**, 286-287.

Engvild KC. 1994 The chloroindole auxins of pea, strong plant growth hormones or endogenous herbicides? Denmark: Risø National Laboratory Denmark.

Engvild KC, Egsgaard H, Larsen E. 1980 Determination of 4‐chloroindole‐3‐acetic acid methyl ester in Lathyrus, Vicia and Pisum by gas chromatography–mass spectrometry. Physiologia Plantarum **48**, 499-503.

Ernstsen A, Sandberg G. 1986 Identification of 4‐chloroindole‐3‐acetic acid and indole‐3‐aldehyde in seeds of *Pinus sylvestris*. Physiologia Plantarum **68**, 511-518.

Fu J, Wang S. 2011 Insights into auxin signaling in plant–pathogen interactions. Frontiers in Plant Science **2**, 74.

Gepts P, Beavis WD, Brummer EC, Shoemaker RC, Stalker HT, Weeden NF, Young ND. 2005 Legumes as a model plant family. Genomics for food and feed report of the cross-legume advances through genomics conference. Plant Physiology **137**, 1228-1235.

Graham PH, Vance CP. 2003 Legumes: importance and constraints to greater use. Plant physiology **131**, 872-877.

Gribble GW. 1998 Naturally occurring organohalogen compounds. Accounts of Chemical Research **31**, 141-152.

Gribble GW. 2010 Naturally Occurring Organohalogen Compounds-A Comprehensive Update: A Comprehensive Update, Vol 91. Vienna: Springer.

Gross JH. 2011 Tandem Mass Spectrometry. *In* Mass Spectrometry. Springer, pp 415-478.

Group LPW. 2013 Legume phylogeny and classification in the 21st century: progress, prospects and lessons for other species-rich clades. Taxon **62**, 217-248.

Hammer PE, Hill DS, Lam ST, Van Pée K-H, Ligon JM. 1997 Four genes from Pseudomonas fluorescens that encode the biosynthesis of pyrrolnitrin. Applied and Environmental Microbiology **63**, 2147-2154.

Hecht H, Sobek H, Haag T, Pfeifer O, Van Pee K-H. 1994 The metal-ion-free oxidoreductase from Streptomyces aureofaciens has an α/β hydrolase fold. Nature Structural and Molecular Biology **1**, 532-537.

Hofmann B, Tölzer S, Pelletier I, Altenbuchner J, Van Pee K, Hecht H. 1998 Structural investigation of the cofactor-free chloroperoxidases. Journal of Molecular Biology **279**, 889-900.

Howieson J, Yates R, Foster K, Real D, Besier R. 2008 Prospects for the future use of legumes. In 'Nitrogen-fixing leguminous symbioses'.(Eds MJ Dilworth, EK James, JI Sprent, WE Newton) pp. 363–393. *In*. Springer, Berlin

Im Kim J, Murphy AS, Baek D, Lee S-W, Yun D-J, Bressan RA, Narasimhan ML. 2011 YUCCA6 over-expression demonstrates auxin function in delaying leaf senescence in Arabidopsis thaliana. Journal of Experimental Botany **62**, 3981-3992.

Jager CE, Symons GM, Nomura T, Yamada Y, Smith JJ, Yamaguchi S, Kamiya Y, Weller JL, Yokota T, Reid JB. 2007 Characterization of two brassinosteroid C-6 oxidase genes in pea. Plant Physiology **143**, 1894-1904.

Johnstone MM, Reinecke DM, Ozga JA. 2005 The auxins IAA and 4-Cl-IAA differentially modify gibberellin action via ethylene response in developing pea fruit. Journal of Plant Growth Regulation **24**, 214-225.

Johnstone MMG, Reinecke DM, Ozga JA. 2005 The auxins IAA and 4-Cl-IAA differentially modify gibberellin action via ethylene response in developing pea fruit. Journal of Plant Growth Regulation **24**, 214-225.

Karcz W, Burdach Z. 2002 A comparison of the effects of IAA and 4 ‐ Cl ‐ IAA on growth, proton secretion and membrane potential in maize coleoptile segments. Journal of Experimental Botany **53**, 1089-1098.

Katayama M, Thiruvikraman SV, Marumo S. 1987 Identification of 4-chloroindole-3-acetic acid and its methyl ester in immature seeds of *Vicia amurensis* (the tribe Vicieae), and their absence from three species of Phaseoleae. Plant and Cell Physiology **28**, 383-386.

Katayama M, Thiruvikraman SV, Marumo S. 1988 Localization of 4-chloroindole-3-acetic acid in seeds of Pisum sativum and its absence from all other organs. Plant and Cell Physiology **29**, 889-891.

Kepinski S, Leyser O. 2005 The Arabidopsis F-box protein TIR1 is an auxin receptor. Nature **435**, 446-451.

Kloda J, Dean P, Maddren C, MacDonald D, Mayes S. 2008 Using principle component analysis to compare genetic diversity across polyploidy levels within plant complexes: an example from British Restharrows (Ononis spinosa and Ononis repens). Heredity **100**, 253-260.

Kodaira K-S, Qin F, Tran L-SP, Maruyama K, Kidokoro S, Fujita Y, Shinozaki K, Yamaguchi-Shinozaki K. 2011 Arabidopsis Cys2/His2 zinc-finger proteins AZF1 and AZF2 negatively regulate abscisic acid-repressive and auxin-inducible genes under abiotic stress conditions. Plant Physiology **157**, 742-756.

Korasick DA, Enders TA, Strader LC. 2013 Auxin biosynthesis and storage forms. Journal of Experimental Botany **64**, 2541-2555.

Lam HK, McAdam S, McAdam E, Ross J. 2015 Evidence that chlorinated auxin is restricted to the Fabaceae but not to the Fabeae. Plant Physiology **168**, 798-803.

Lavin M, Doyle JJ, Palmer JD. 1990 Evolutionary significance of the loss of the chloroplast-DNA inverted repeat in the Leguminosae subfamily Papilionoideae. Evolution **44**, 390-402.

Lavin M, Herendeen PS, Wojciechowski MF. 2005 Evolutionary rates analysis of Leguminosae implicates a rapid diversification of lineages during the Tertiary. Systematic Biology **54**, 575-594.

Leblanc C, Vilter H, Fournier J-B, Delage L, Potin P, Rebuffet E, Michel G, Solari P, Feiters M, Czjzek M. 2015 Vanadium haloperoxidases: From the discovery 30 years ago to X-ray crystallographic and V K-edge absorption spectroscopic studies. Coordination Chemistry Reviews **301**, 134-146.

Ludwig-Müller J. 2011 Auxin conjugates: their role for plant development and in the evolution of land plants. Journal of Experimental Botany **62**, 1757-1773.

Mano Y, Nemoto K. 2012 The pathway of auxin biosynthesis in plants. Journal of Experimental Botany **63**, 2853-2872.

Martin E, Akan H, Ekici M, Aytac Z. 2011 Karyotype analyses of ten sections of Trigonella (Fabaceae). Comparative cytogenetics **5**, 105-121.

McAdam EL, Meitzel T, Quittenden LJ, Davidson SE, Dalmais M, Bendahmane AI, Thompson R, Smith J, Nichols D, Urquhart S, Gélinas-Marion A, Aubert G, Ross J. 2017 Evidence that Auxin is Required for Normal Seed Size and Starch Synthesis in Pea. New Phytologist **In Press**

Mikić A, Smýkal P, Kenicer G, Vishnyakova M, Sarukhanyan N, Akopian J, Vanyan A, Gabrielyan I, Smýkalová I, Sherbakova E. 2013 The bicentenary of the research on 'beautiful'vavilovia (Vavilovia formosa), a legume crop wild relative with taxonomic and agronomic potential. Botanical Journal of the Linnean Society **172**, 524-531.

Mueller-Roeber B, Balazadeh S. 2013 Auxin and Its Role in Plant Senescence. Journal of Plant Growth Regulation **33**, 1-13.

Murphy C. 2003 New frontiers in biological halogenation. Journal of Applied Microbiology **94**, 539-548.

Mutka AM, Fawley S, Tsao T, Kunkel BN. 2013 Auxin promotes susceptibility to Pseudomonas syringae via a mechanism independent of suppression of salicylic acid-mediated defenses. The Plant Journal **74**, 746-754.

Neff MM, Turk E, Kalishman M. 2002 Web-based primer design for single nucleotide polymorphism analysis. TRENDS in Genetics **18**, 613-615.

Neumann CS, Fujimori DG, Walsh CT. 2008 Halogenation strategies in natural product biosynthesis. Chemistry & Biology **15**, 99-109.

Nicholas K, Nicholas H. 1997 GeneDoc: a tool for editing and annotating multiple sequence alignments. *In*, Vol Ed. GeneDoc

Nichols P, Revell C, Humphries A, Howie J, Hall E, Sandral G, Ghamkhar K, Harris C. 2012 Temperate pasture legumes in Australia—their history, current use, and future prospects. Crop and Pasture Science **63**, 691-725.

Obroucheva N. 2014 Hormonal regulation during plant fruit development. Russian Journal of Developmental Biology **45**, 11-21.

Ohshiro T, Littlechild J, Garcia-Rodriguez E, Isupov MN, Iida Y, Kobayashi T, Izumi Y. 2004 Modification of halogen specificity of a vanadium-dependent bromoperoxidase. Protein Science **13**, 1566-1571.

Oskoueiyan R, Osaloo SK, Maassoumi AA, Nejadsattari T, Mozaffarian V. 2010 Phylogenetic status of Vavilovia formosa (Fabaceae-Fabeae) based on nrDNA ITS and cpDNA sequences. Biochemical Systematics and Ecology **38**, 313-319.

Ozga JA, Reinecke DM, Ayele BT, Ngo P, Nadeau C, Wickramarathna AD. 2009 Developmental and hormonal regulation of gibberellin biosynthesis and catabolism in pea fruit. Plant Physiology **150**, 448-462.

Park S, Ozga JA, Cohen JD, Reinecke DM. 2010 Evidence of 4-Cl-IAA and IAA bound to proteins in pea fruit and seeds. Journal of Plant Growth Regulation **29**, 184-193.

Parry G, Calderon-Villalobos L, Prigge M, Peret B, Dharmasiri S, Itoh H, Lechner E, Gray W, Bennett M, Estelle M. 2009 Complex regulation of the TIR1/AFB family of auxin receptors. Proceedings of the National Academy of Sciences **106**, 22540-22545.

Pattison RJ, Csukasi F, Catalá C. 2014 Mechanisms regulating auxin action during fruit development. Physiologia Plantarum **151**, 62-72.

Peters WS, Haffer D, Hanakam CB, van Bel AJ, Knoblauch M. 2010 Legume phylogeny and the evolution of a unique contractile apparatus that regulates phloem transport. American Journal of Botany **97**, 797-808.

Petropoulos GA. 2002 Fenugreek: the genus Trigonella, CRC Press.

Pless T, Böttger M, Hedden P, Graebe J. 1984 Occurrence of 4-Cl-indoleacetic acid in broad beans and correlation of its levels with seed development. Plant Physiology **74**, 320-323.

Quittenden LJ, Davies NW, Smith JA, Molesworth PP, Tivendale ND, Ross JJ. 2009 Auxin biosynthesis in pea: characterization of the tryptamine pathway. Plant Physiology **151**, 1130-1138.

Ranjbar M, Hajmoradi Z. 2015 A new species of Trigonella sect. Ellipticae (Leguminosae-Papilionoideae) from Iran, including cytogenetic and anatomical notes. Phytotaxa **202**, 26-34.

Ranjbar M, Zahra H. 2016 Chromosome numbers and biogeography of the genus Trigonella (Fabaceae). Caryologia **69**, 223-234.

Reinecke DM. 1999 4-Chloroindole-3-acetic acid and plant growth. Plant Growth Regulation **27**, 3-13.

Reinecke DM, Ozga JA, Ilić N, Magnus V. 1999 Molecular properties of 4-substituted indole-3-acetic acids affecting pea pericarp elongation. Plant Growth Regulation **27**, 39-48.

Reinecke DM, Ozga JA, Magnus V. 1995 Effect of halogen substitution of indole-3-acetic acid on biological activity in pea fruit. Phytochemistry **40**, 1361-1366.

Rescher U, Walther A, Schiebl C, Klämbt D. 1996 In vitro binding affinities of 4-chloro-, 2-methyl-, 4-methyl-, and 4-ethylindoleacetic acid to auxin-binding protein 1 (ABP1) correlate with their growth-stimulating activities. Journal of Plant Growth Regulation **15**, 1-3.

Rose TM, Schultz ER, Henikoff JG, Pietrokovski S, McCallum CM, Henikoff S. 1998 Consensus-degenerate hybrid oligonucleotide primers for amplification of distantly related sequences. Nucleic Acids Research **26**, 1628-1635.

Ross JJ, Reid JB. 2010 Evolution of growth-promoting plant hormones. Functional Plant Biology **37**, 795-805.

Ross JJ, Tivendale ND, Davidson SE, Reid JB, Davies NW, Quittenden LJ, Smith JA. 2012 A mutation affecting the synthesis of 4-chloroindole-3-acetic acid. Plant Signaling and Behavior **7**, 1533-1536.

Runguphan W, Qu X, O'Connor SE. 2010 Integrating carbon-halogen bond formation into medicinal plant metabolism. Nature **468**, 461-464.

Sauer M, Robert S, Kleine-Vehn J. 2013 Auxin: simply complicated. Journal of Experimental Botany **64**, 2565-2577.

Schaefer H, Hechenleitner P, Santos-Guerra A, de Sequeira MM, Pennington RT, Kenicer G, Carine MA. 2012 Systematics, biogeography, and character evolution of the legume

tribe Fabeae with special focus on the middle-Atlantic island lineages. BMC Evolutionary Biology **12**, 250.

Serrani JC, Ruiz‐Rivero O, Fos M, García‐Martínez JL. 2008 Auxin‐induced fruit‐set in tomato is mediated in part by gibberellins. The Plant Journal **56**, 922-934.

Shaw PD, Hager LP. 1959 An enzymatic chlorination reaction. Journal of the American Chemical Society **81**, 1011-1012.

Singh A. 1973 Studies on the interspecific hybrids of Trigonella corniculata L., T. hamosa L. and T. cretica L. Genetica **44**, 264-269.

Simon S, Petrášek J. 2011 Why plants need more than one type of auxin. Plant Science **180**, 454-460.

Sirjaev G. 1928-1934 Generis Trigonella L. Revisio critica. Brno: Faculty of Science of the University of Masary.

Small E, Lassen P, Brookes BS. 1987 An expanded circumscription of Medicago (Leguminosae, Trifolieae) based on explosive flower tripping. Willdenowia **16**, 415-437.

Smýkal P, Coyne CJ, Ambrose MJ, Maxted N, Schaefer H, Blair MW, Berger J, Greene SL, Nelson MN, Besharat N. 2015 Legume crops phylogeny and genetic diversity for science and breeding. Critical Reviews in Plant Sciences **34**, 43-104.

Steele KP, Ickert-Bond SM, Zarre S, Wojciechowski MF. 2010 Phylogeny and character evolution in Medicago (Leguminosae): Evidence from analyses of plastid trnK/matK and nuclear GA3ox1 sequences. American Journal of Botany **97**, 1142-1155.

Steele KP, Wojciechowski MF. 2003 Phylogenetic analyses of tribes Trifolieae and Vicieae, based on sequences of the plastid gene matK (Papilionoideae: Leguminosae). *In* B Kilitgaard, A Bruneau, eds, Advances in Legume Systematics, part 10. Royal Botanical Gardens, Kew, pp 355-370.

Steffens B, Lüthen H. 2000 New methods to analyse auxin-induced growth II: The swelling reaction of protoplasts–a model system for the analysis of auxin signal transduction? Plant Growth Regulation **32**, 115-122.

Symons GM, Ross JJ, Murfet IC. 2002 The bushy pea mutant is IAA‐deficient. Physiologia Plantarum **116**, 389-397.

Tang N, Deng W, Hu G, Hu N, Li Z. 2015 Transcriptome profiling reveals the regulatory mechanism underlying pollination dependent and parthenocarpic fruit set mainly mediated by auxin and gibberellin. PloS one **10**, e0125355.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. 1997 The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic acids research **25**, 4876-4882.

Tivendale ND, Davidson SE, Davies NW, Smith JA, Dalmais M, Bendahmane AI, Quittenden LJ, Sutton L, Bala RK, Le Signor C. 2012 Biosynthesis of the halogenated auxin, 4-chloroindole-3-acetic acid. Plant Physiology **159**, 1055-1063.

Tivendale ND, Davies NW, Molesworth PP, Davidson SE, Smith JA, Lowe EK, Reid JB, Ross JJ. 2010 Reassessing the role of *N*-hydroxytryptamine in auxin biosynthesis. Plant Physiology **154**, 1957-1965.

Tivendale ND, Ross JJ, Cohen JD. 2014 The shifting paradigms of auxin biosynthesis. Trends in Plant Science **19**, 44-51.

Unversucht S, Hollmann F, Schmid A, van Pée KH. 2005 FADH2-dependence of tryptophan 7-halogenase. Advanced Synthesis & Catalysis **347**, 1163-1167.

Vaillancourt FH, Yeh E, Vosburg DA, Garneau-Tsodikova S, Walsh CT. 2006 Nature's inventory of halogenation catalysts: Oxidative strategies predominate. Chemical Reviews **106**, 3364-3378.

Van Berkel W, Kamerbeek N, Fraaije M. 2006 Flavoprotein monooxygenases, a diverse class of oxidative biocatalysts. Journal of biotechnology **124**, 670-689.

Van Pée K-H. 2012 Enzymatic chlorination and bromination. Methods in Enzymology **516**, 237-257.

van Pée K-H, Zehner S. 2003 Enzymology and molecular genetics of biological halogenation. *In* Natural production of organohalogen compounds. Springer, pp 171-199.

van Pée K, Ludwig-Müller J. 2002 Halogenated indole derivatives. Current Topics in Phytochemistry **5**, 1-21.

Vanneste S, Friml J. 2009 Auxin: a trigger for change in plant development. Cell **136**, 1005-1016.

Wagner C, El Omari M, König GM. 2009 Biohalogenation: nature's way to synthesize halogenated metabolites. Journal of Natural Products **72**, 540-553.

Wang S, Bai Y, Shen C, Wu Y, Zhang S, Jiang D, Guilfoyle TJ, Chen M, Qi Y. 2010 Auxin-related gene families in abiotic stress response in Sorghum bicolor. Functional & Integrative Genomics **10**, 533-546.

Weichold V, Milbredt D, van Pée KH. 2016 Specific Enzymatic Halogenation—From the Discovery of Halogenated Enzymes to Their Applications In Vitro and In Vivo. Angewandte Chemie International Edition **55**, 6374-6389.

Wojciechowski MF, Lavin M, Sanderson MJ. 2004 A phylogeny of legumes (Leguminosae) based on analysis of the plastid *matK* gene resolves many well-supported subclades within the family. American Journal of Botany **91**, 1846-1862.

Wojciechowski MF, Sanderson MJ, Steele KP, Liston A. 2000 Molecular phylogeny of the "temperate herbaceous tribes" of papilionoid legumes: a supertree approach. Advances in Legume Systematics **9**, 277-298.

Wong C, Fujimori DG, Walsh CT, Drennan CL. 2009 Structural analysis of an open active site conformation of nonheme iron halogenase CytC3. Journal of the American Chemical Society **131**, 4872-4879.

Woodward AW, Bartel B. 2005 Auxin: regulation, action, and interaction. Annals of Botany **95**, 707-735.

Xu G, Wang B-G. 2016 Independent Evolution of Six Families of Halogenating Enzymes. PloS one **11**, e0154619.

Yeh E, Garneau S, Walsh CT. 2005 Robust in vitro activity of RebF and RebH, a two-component reductase/halogenase, generating 7-chlorotryptophan during rebeccamycin biosynthesis. Proceedings of the National Academy of Sciences of the United States of America **102**, 3960-3965.

Young ND, Udvardi M. 2009 Translating Medicago truncatula genomics to crop legumes. Current Opinion in Plant Biology **12**, 193-201.

Zehner S, Kotzsch A, Bister B, Süssmuth RD, Méndez C, Salas JA, van Pée K-H. 2005 A regioselective tryptophan 5-halogenase is involved in pyrroindomycin biosynthesis in Streptomyces rugosporus LL-42D005. Chemistry and Biology **12**, 445-452.

Zhang S, van Duijn B. 2014 Cellular auxin transport in algae. Plants **3**, 58-69.

Appendix 1: Primer details

Gene	Purpose	Primer names	Primer sequences (5'-3')	Tm(°C)
ACT	qRT-PCR	ACT-F	GTGTCTGGATTGGAGGATCAATC	59
		ACT-R	GGCCAGGCTCATCATATTCA	56
PsVHPO1	Genotyping for *Psvhpo1*by dCAPs marker	PsHPO1-2760-DraI-F	TTTTTCCAGGTTCTTAATCAAATTGTATTT	55
		VHPO1-HRM-2R	CTCGGATCCATTAAACCATCA	57
	Gene isolation	PsVHPO1B1265gF	ATGCCGTCTTCTCATTCAGC	52
		VHPO1ex34R1	TACTTTAGGTTGCCGAAGAA	48
		PsVHPO1778R	TTACTTGGAATCGTAACAGCATCTA	53
		PsVHPO1635F	AACACTAGAGACTTGTTGTTTTGGA	53
	Genotyping by sequencing	VHPO1ex34F1	GTGGCAGGTGATGTATGATG	52
		VHPO1ex34R1	TACTTTAGGTTGCCGAAGAA	48
	HRM (Line/Allele 2760)	VHPO1-HRM-1F	AATGACCGGGTTTTTCCAG	55
		VHPO1-HRM-1R	CGCGAAGAGGTCTGCTGTCA	62
	HRM (Line/Allele 3023)	VHPO1-HRM-2F	TCTGACAGCAGACCTCTTCG	60
		VHPO1-HRM-2R	CTCGGATCCATTAAACCATCA	57
	qRT-PCR	PsVHPO1F1	TGTATTTATAACCTTTCCTTTTCACC	52
		PsVHPO1R1	CCATAAGACAGACGCTTCCA	52
PsVHPO2	Gene isolation	PsVHPO2758F	GCAAGCATTCAAGCGAGTAAAG	53
		PsVHPO2858R	ATACACAAGAAATCTGAATGTCCAAA	52
	qRT-PCR	PsVHPO2F1	TGCCAGAGAAAGAGAAGGAGA	52
		PsVHPO2R1	TTGTGTTCAATCATGTGTTCAATC	50
PsVHPO3	Gene isolation	PsVHPO3386F	ACGCCAAGGTCAAGATTAGTCC	55
		PsVHPO3470R	GCCACAAACCATGCAAACAAA	50
	qRT-PCR	PsVHPO3F1	TTGGTGTTACGACGCTCTTC	52
		PsVHPO3R1	ATCTTCATTTACAGAATACTCATGGTG	54